

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 11:45:25 ; Search time 25.1 Seconds
(without alignments)
1959.548 Million cell updates/sec

Title: US-09-590-991-6

Perfect score: 3454

Sequence: 1 YPVVLADTSSEDLNISDK.....PGKNIKSTLVNCKSTFGY 664

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq.1101.*
1: /SID8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3454	100.0	664	22	AA1980.DAT.*
2	3454	100.0	2120	21	AA1981.DAT.*
3	3269	94.6	1007	19	AA1982.DAT.*
4	803	23.2	192	19	AA1983.DAT.*
5	433	13.1	1946	19	AA1984.DAT.*
6	447.5	13.0	1962	12	AA1985.DAT.*
7	440.5	12.8	1962	12	AA1986.DAT.*
8	438.5	12.7	1962	12	AA1987.DAT.*
9	435.5	12.6	1968	12	AA1988.DAT.*
10	435	12.6	1959	12	AA1989.DAT.*
11	434.5	12.6	1962	12	AA1990.DAT.*

12	434.5	12.6	1962	12	AA1991.DAT.*
13	434.5	12.6	1962	12	AA1992.DAT.*
14	431.5	12.5	1974	12	AA1993.DAT.*
15	429.5	12.4	1560	10	AA1994.DAT.*
16	343.5	9.9	806	13	AA1995.DAT.*
17	302	8.7	1164	18	AA1996.DAT.*
18	302	8.7	1164	21	AA1997.DAT.*
19	298	8.6	1150	18	AA1998.DAT.*
20	296.5	8.6	1150	21	AA1999.DAT.*
21	293	8.5	1167	18	AA2000.DAT.*
22	293	8.5	1167	21	AA2001.DAT.*
23	292	8.5	1181	21	AA2002.DAT.*
24	256	7.4	731	15	AA2003.DAT.*
25	240.5	7.0	690	11	AA2004.DAT.*
26	237.5	6.9	690	18	AA2005.DAT.*
27	237.5	6.9	690	18	AA2006.DAT.*
28	236	6.8	654	18	AA2007.DAT.*
29	236	6.8	654	20	AA2008.DAT.*
30	229.5	6.6	659	18	AA2009.DAT.*
31	229.5	6.6	659	20	AA2010.DAT.*
32	214.5	6.2	659	18	AA2011.DAT.*
33	204	5.9	766	21	AA2012.DAT.*
34	201.5	5.8	522	18	AA2013.DAT.*
35	201	5.8	522	20	AA2014.DAT.*
36	201	5.8	522	20	AA2015.DAT.*
37	199.5	5.7	382	17	AA2016.DAT.*
38	196	5.6	1316	21	AA2017.DAT.*
39	194	5.6	1337	21	AA2018.DAT.*
40	194	5.6	748	21	AA2019.DAT.*
41	192	5.6	748	21	AA2020.DAT.*
42	188.5	5.5	769	21	AA2021.DAT.*
43	187.5	5.4	381	8	AA2022.DAT.*
44	187.5	5.4	829	21	AA2023.DAT.*
45	187.5	5.4	829	21	AA2024.DAT.*

ALIGNMENTS

RESULT 1	AA1980.DAT.*
ID	AA1980.DAT.*
AC	AA1980.DAT.*
XX	AA1980.DAT.*
XX	AA1980.DAT.*
DE	20-APR-2001 (first entry)
XX	S. pneumoniae Sp128 polypeptide.
XX	Immunogenic: Sp128; Sp130; pneumococcal; otitis media; nasopharyngeal;
KW	bronchial; Lung; blood; Infection; Immune response; Immunotherapy;
KW	antibacterial; auditory; vaccine.
OS	Streptococcus pneumoniae.
XX	WO200076540-A2.
PN	21-DEC-2000.
XX	09-JUN-2000; 2000WO-US15925.
PF	10-JUN-1999; 99US-0138453.
PR	(MEDT-) MED IMMUNE INC.
PA	Adamou JE, Choi GH;
XX	WPI: 2001-112197/12.
DR	N-PSDB; AAC84741.
XX	New vaccines comprising Sp128 or Sp130 polypeptides, for treating and
PT	preventing pneumococcal infections, particularly infections caused by
PT	Streptococcus, e.g. otitis media, nasopharyngeal, bronchial, lung or

Mutant protease (N
Mutant protease (K
Mutant protease (d
S. cremoris protei
Rp-III residual pr
Streptococcal C4a
SCP4a3 peptidase
Streptococcal C4a
SCP8 peptidase (w1
Streptococcal C4a
SCP12 peptidase
SCP1 peptidase (w
Melon protease fro
Residual protease-
Bacillus subtilis
Residual protease
Pyrococcus furiosu
Hyperthermostable
Thermococcus prote
MO9856926 Seg ID 1
Protease. Synthet
Subtilase (St_P2SC
MprA protein. Bac
Pyrococcus furiosu
Hyperthermostable
Hyperthermostable
Subtilisin N62D/Y1
A calcium-dependen
Arabidopsis thaliana
Arabidopsis thaliana
Bacillus carlsberg
Subtilisin
Subtilase (Pigesp)

PT blood infections -
 XX
 PS Claim 4; Page 47-50; 54pp: English.
 XX
 CC The invention relates to novel immunogenic polypeptides, Spi28 and Spi30
 CC from *S. pneumoniae*. Vaccines comprising the polypeptides are useful for
 CC the treatment and prevention of pneumococcal infections, particularly
 CC infections caused by *Streptococcus*, such as otitis media, nasopharyngeal,
 CC bronchial, lung or blood infections. The antigens are used as immunogenic
 CC agents to stimulate an immune response. The antisera and antibodies may
 CC also be used in diagnosing and treating pneumococcal infections.
 CC Recombinant polypeptides serve as a mechanism for stimulating production
 CC of antibodies for use in passive immunotherapy, diagnostic reagents, and
 CC as reagents in other processes such as affinity chromatography. The
 CC present sequence represents the *S. pneumoniae* Spi28 polypeptide.
 XX
 SQ Sequence 664 AA:

Query Match 100.0%; Score 3454; DB 22; Length 664;
 Best Local Similarity 100.0%; Pred. No. 6, 8e-195;
 Matches 664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YPVVLADTSSSEDALNISDEKVAENKEHENIHSAMETSDQFEKKTAVIKEEVVSKN 60
 DB 1 YPVVLADTSSSEDALNISDEKVAENKEHENIHSAMETSDQFEKKTAVIKEEVVSKN 60
 QY 61 PYIDNNTSNEAKIKKEENSNGSGDYTDSPVKNKTENPKKEDKVVYIAEFKDSKESGKAI 120
 DB 61 PYIDNNTSNEAKIKKEENSNGSGDYTDSPVKNKTENPKKEDKVVYIAEFKDSKESGKAI 120
 QY 121 KELSLKNTFVLYTYDRIFNGSAIETTPDNLDKIKOIGISSVERAKQVOPMNHARKEI 180
 DB 121 KELSLKNTFVLYTYDRIFNGSAIETTPDNLDKIKOIGISSVERAKQVOPMNHARKEI 180
 QY 181 GVEBAIDYLSINAPESCKNFDGKGMVTSNIDTGYRHKAMRIDDAKASREKKEDIKG 240
 DB 181 GVEBAIDYLSINAPESCKNFDGKGMVTSNIDTGYRHKAMRIDDAKASREKKEDIKG 240
 QY 241 TDKNVWISDKIRPAFNYYNGKIVKDYDGRDYPDPHGMIACTILAGNDPDIKKNNG 300
 DB 241 TDKNVWISDKIRPAFNYYNGKIVKDYDGRDYPDPHGMIACTILAGNDPDIKKNNG 300
 QY 301 IDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIDSIKHNVDVYSSSGFTGTGLVGERKY 360
 DB 301 IDGIAPNAQIFSYKMYSDAGSGFAGDETMFHAIDSIKHNVDVYSSSGFTGTGLVGERKY 360
 QY 361 WQAIRALKRKGIPMYVATGNTATSSASSSMDLVANHLKMTDTGNTVTTAAHEDAIYAS 420
 DB 361 WQAIRALKRKGIPMYVATGNTATSSASSSMDLVANHLKMTDTGNTVTTAAHEDAIYAS 420
 QY 421 AKKQTVFEDKYNIGSEFFKRNIGAFDKSKITTNEDGTAKPSKLFYIKGDDODLIG 480
 DB 421 AKKQTVFEDKYNIGSEFFKRNIGAFDKSKITTNEDGTAKPSKLFYIKGDDODLIG 480
 QY 481 LDLRGKIAVMDRIYTKDLKNAFKKAMDGARAIWVNTVYVNRDNTETLRPAKYEADG 540
 DB 481 LDLRGKIAVMDRIYTKDLKNAFKKAMDGARAIWVNTVYVNRDNTETLRPAKYEADG 540
 QY 541 TKSGVFESISGDDGKLMNMNINPKKTEVKRRNKEDFKDLQOYPIIDMESNSKPNVGD 600
 DB 541 TKSGVFESISGDDGKLMNMNINPKKTEVKRRNKEDFKDLQOYPIIDMESNSKPNVGD 600
 QY 601 EKEDIFKAPDTDEKELKEDIIVPAGSTSMGPRIDLKLPVSAFGKNIKSTLVNTJNKS 660
 DB 601 EKEDIFKAPDTDEKELKEDIIVPAGSTSMGPRIDLKLPVSAFGKNIKSTLVNTJNKS 660
 QY 661 TYGY 664
 DB 661 TYGY 664

RESULT 2

AAV81710
 ID AAV81710 standard; Protein; 2120 AA.
 XX
 AC AAV81710;
 XX
 DT 02-JUN-2000 (first entry)
 XX
 DE Streptococcus pneumoniae protein sequence ID3.
 XX
 KW Streptococcus pneumoniae infection; immunogen; antigen; diagnosis; AIDS;
 KW bacterial pneumonia; asplenia; heart disease; lung disease; alcoholism;
 KW kidney disease; diabetes; immunosuppressive disorder; otitis media;
 KW pneumococcal septicemia; sinusitis; meningitis; therapy.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN W0200006738-A2.
 XX
 PD 10-FEB-2000.
 XX
 PE 27-JUL-1999; 99WO-GB02452.
 XX
 PR 27-JUL-1998; 98GB-0016336.
 XX
 PR 19-MAR-1999; 99US-0125329.
 XX
 PA (MICR-) MICROBIAL TECHNICS LTD.
 XX
 PI Le Page RWF, Wells JM, Haniffy SB, Hansbro PM;
 XX
 DR WPI: 2000-195301/17.
 XX
 DR N-PSDB: AA291806.
 XX
 PS Claim 2; Page 41-42; 76pp: English.
 XX
 SS
 CC This sequence represents a streptococcus pneumoniae protein of the
 CC invention. The proteins (or their homologues, derivatives and/or
 CC fragments) are useful as immunogens or antigens. Immunogenic or antigenic
 CC compositions comprising the proteins are useful as vaccines and also in
 CC diagnostic assays. The sequences are useful for the detection or
 CC diagnosis of *S. pneumoniae* infection, by contacting a sample to be tested
 CC with them. Agents capable of antagonising, inhibiting or interfering with
 CC the function or expression of the protein or polypeptide are useful in
 CC medical compositions in the treatment or prophylaxis of *S. pneumoniae*
 CC infection. As the sequences can be used to treat *S. pneumoniae* infection,
 CC they can be used to treat bacterial pneumonia, which has high rates in
 CC young children, the elderly, and in patients with predisposing conditions
 CC such as asplenia, heart, lung and kidney disease, diabetes, alcoholism,
 CC or with immunosuppressive disorders, especially AIDS. They can also be
 CC used to treat pneumococcal septicemia, otitis media, sinusitis, and
 CC meningitis.
 CC
 XX
 SQ Sequence 2120 AA:

Query Match 100.0%; Score 3454; DB 21; Length 2120;
 Best Local Similarity 100.0%; Pred. No. 2, 9e-194;
 Matches 664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YPVVLADTSSSEDALNISDEKVAENKEHENIHSAMETSDQFEKKTAVIKEEVVSKN 60
 DB 1 YPVVLADTSSSEDALNISDEKVAENKEHENIHSAMETSDQFEKKTAVIKEEVVSKN 60
 QY 61 PYIDNNTSNEAKIKKEENSNGSGDYTDSPVKNKTENPKKEDKVVYIAEFKDSKESGKAI 120
 DB 61 PYIDNNTSNEAKIKKEENSNGSGDYTDSPVKNKTENPKKEDKVVYIAEFKDSKESGKAI 120
 QY 121 KELSLKNTFVLYTYDRIFNGSAIETTPDNLDKIKOIGISSVERAKQVOPMNHARKEI 180
 DB 121 KELSLKNTFVLYTYDRIFNGSAIETTPDNLDKIKOIGISSVERAKQVOPMNHARKEI 180

```

QY 181 GVEAIDLYKSTNAPFGKNGFNGMYISNIDTGYRHKAMRIDDDAKASMRKEDLK 240
    |||||||
DB 181 yveeaIdYlkslnapfGknfdgrgwnlsnIdtgldYrHkamrIdddakasmrFkEdlkY 240
QY 241 TDKNYLSKIPAFNYVNGKRTTEVKYDDGRFPDPHGHTAGIAGNDETDIKNPN 300
    |||||||
DB 241 tdknylskIphafnyngkrttevkYddgrfDpHgmhlaGlaGndtdIkNfng 300
QY 301 IDGIAPNAOIFSYKMYSDAGSGFAGDETFMHAIEDSIKHNVDVSVSSGFTGVLGEXY 360
    |||||||
DB 301 IdgIapnaqlfSykmysdagsgfagdeTFmhaIedslkhnvdvsvsgftgvlGexky 360
QY 361 WQAIRLRKAGIPMYVATGNATSSASSSDLVANNHLMKMTDGTNVRTIAHEDAIAS 420
    |||||||
DB 361 wqairlrkagIpmvvaTgnatssassswdlvannhlmkmtDgnvrtIaahedaIavas 420
QY 421 AKKQTEPFPKVNITGSESPFRNIGAFEDSKITNEDGTRKAPSKLFEVYIGKODDGLG 480
    |||||||
DB 421 akqtefEdvniTgsefkyrniGafEdkskItnedgtrkapskLfvyIgKqddqdlIy 480
QY 481 LDRGKIAVMDRIYTRDLKNAFKKAMDGCARAIMVTVNVYVRDMMWTELPAMGYEAD 540
    |||||||
DB 481 ldrGkIavmdRIytrDlKnaFkKamDgcaraImvtnvYvrDmmwTelPamgyeadeG 540
QY 541 TKSQVPSISGDSGVKLMNMNINPDKTEVKRNKEDFKLEOYYPIDMESFNSNKNPVG 600
    |||||||
DB 541 tksqvpsIsGdsgvKlMnmNinpdktevrnkEdfkKleoyYpIdmesfnsnKnpvgd 600
QY 601 EKRIDKFPADPKELKEDITYPAGSTSGRPRIIDLLKRDVSAPKNIKSTLVNNGKS 660
    |||||||
DB 601 eKridkfpadPkElkEdiTyPaGstsgRpriIdllKrdvsApknIKstlvnngks 660
QY 661 TTYG 664
    ||||
DB 661 tygy 664

RESULT 3
AAM61246
ID AAM61246 standard: Protein; 1007 AA.
XX
AC AAM61246;
XX
DE 02-OCT-1998 (first entry)
XX
DE Streptococcus pneumoniae SPI22 protein.
XX
DE Streptococcus pneumoniae: antigen; vaccine; infection; diagnosis;
detection; pneumonia; otitis media; meningitis.
XX
OS Streptococcus pneumoniae.
XX
PN MO9818930-A2.
XX
PD 07-MAY-1998.
XX
PF 30-OCT-1997; 97MO-US19422.
XX
PR 31-OCT-1996; 96US-0029960.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Chol GH, Hromockyj A, Johnson LS, Kunsch CA;
XX
DR MPI: 1998-272224/24.
XX
DR N-PSDB; AAV27431.
XX
PT Nucleic acid encoding antigenic peptide(s) from Streptococcus
PT pneumoniae - or their epitope-containing fragments, useful in
PT protective or therapeutic vaccines, and for diagnosis
XX
PS Claim 11; Page 92-93; 118pp; English.
XX

```

```

CC The present sequence represents a protein from Streptococcus pneumoniae.
CC The nucleic acid sequence encoding the Streptococcus pneumoniae protein
CC can be useful in vaccines for inducing protective antibodies against
CC Streptococcus pneumoniae, for treatment or prevention of infection e.g.
CC pneumonia, otitis media or meningitis. Probes based on the nucleic acid
CC are used to detect Streptococcus infection (by usual hybridisation or
CC amplification methods), also for isolating Streptococcus genes or their
CC allelic variants. The protein can be used similarly to detect specific
CC antibodies in standard immunoassays, especially for diagnosing or
CC monitoring infections. Antibodies which bind the protein are used to
CC detect corresponding antigens, to purify the protein and for passive
CC immunisation (optionally coupled to a toxin). Vaccines are administered,
CC e.g. by injection, orally or through the skin. Vaccines are administered,
CC (especially 10-300) mu g/ml per dose.
XX
SQ Sequence 1007 AA;
QY 38 ETSQDEKREKKTAVIKREYVSKNPVIDNNTSNEAKIKKESNKSOGDYTDSPVNKNTEN 97
    |||||||
DB 1 etsqdIkEkktavIkEkevskNpviDnntSneakIkEesnksSgdytdspvnknten 60
QY 98 PKKEDKVYVIAEERDKESGEKAIKELSLKNTKVLTYTPRIENGSAIETTPMDIKIKOI 157
    |||||||
DB 61 pKkedkvYvIaeERdkEsgekaIkElslkntKvltYtpriEngsaIettpmdIkIkdI 120
QY 158 EGISSEVERAKQVOPMMNHARKEIGVEAIDLYKSNAPFGKNGFNGMYISNIDGTDR 217
    |||||||
DB 121 egIsseveraqkvOpmmnharkeIgveaIdYlksInapfgknGfngmyIsnIdgtDyr 180
QY 218 HKAMRIDDDAKASMRKEDLKCTOKNMYLSDKIRPAFNYNGKRTTEVKYDDGRFP 277
    |||||||
DB 181 hKamrIdddakasmrFkEdlkCtOknyvlsdkIrpafnyngkrttevkYddgrfDp 240
QY 278 HGMHIGILAGNDETDIKNFGIDCIAPNAQIFSYKMYSDAGSGFAGDETFMHAIEDSI 337
    |||||||
DB 241 hgmhIgilagndtdIkNfngIdcIapnaQIfsykmysdagsgfagdeTFmhaIedsl 300
QY 338 KHNVDVSVSSGFTGVLGEXYQWAIKALRKAGIPMYVATGNATSSASSSDLVANNH 397
    |||||||
DB 301 khnvDvsvsgftgvlGexyQwAIrAlrkagIpmvvaTgnatssassswdlvannh 360
QY 398 LKMTDGTNVRTIAHEDAIASAKNQVFEKVNITGSESPFRNIGAFEDSKITNED 457
    |||||||
DB 361 lKmtDgnvrtIaahedaIavasaknQvfeKvniTgsefkyrniGafEdkskItned 420
QY 458 GTRKAPSKLFEVYIGKODDGLIGLRGKIAVMDRIYTRDLKNAFKKAMDGCARAIMVN 517
    |||||||
DB 421 gtrkapskLfvyIgKqddqdlIglrGkIavmdRIytrDlKnaFkKamDgcaraImvN 480
QY 518 TVNVYVRDMMWTELPAMGYEADGKTSQVPSISGDSGVKLMNMNINPDKTEVKRNKEDFK 577
    |||||||
DB 481 tvnvYvrDmmwTelPamgyeadeGtksqvpsIsGdsgvKlMnmNinpdktevrnkEdfk 540
QY 578 DKLEOYYPIDMESFNSNKNPVGDEKEDIFKAPDTPKELKEDITYPAGSTSGRPRIID 637
    |||||||
DB 541 dKleoyYpIdmesfnsnKnpvgdeKedIfkApdtpKElkEdiTyPaGstsgRpriIdlI 600
QY 638 LKPDYSAFGKNIKSTLVNNGKSTGYG 664
    |||||||
DB 601 lKpdysaPgknIKstlvnngkStgyg 627

RESULT 4
AAY86159
ID AAY86159 standard: Protein; 192 AA.
XX
AC AAY86159;
XX

```


Oy		408	RTAAHEADIAVASAKNOTVEEDKV--NIGESPKRYRN--TGAFEDSKITTNEDDGTKA--	466
Dd		412	-----pdaliivaaseemskvltldvcdeilgvtlssnselkga---aqvltqlsenysvl	462
Oy		462	PSKLFYVYGKGQODDIDG---IDLRGKIIVMDR-IYTNDLKNAPFKKAMDGCARAIIMVN	517
Dd		463	tkklklkvdmjggsgddytaekkeevkgqglavkrqaytfsakvanaka--agaagivlyln	520
Oy		518	TWNYYNRDNMTLEPAMGTEADEGTSKOVFSISGDGCVKLMMINIPDKTKTEVRNNKKDPK	577
Dd		521	-----seddg--llsmajldcktfprljgmksadqk	547
Oy		578	DKLEOYVPIDMESFNNSKNPNVGDE---KEIDKFAPPTOKELKEDIITYPAOSTSMGPRI	634
Dd		548	fwlqpkqvkkvaasrlkfctalidhsragkmsdf-----tswgpqp	586
Oy		635	DLLLKPDVSAPFGKNIKSTLN	654
		587	: : eldfxfpeltapqgklyalan	606
		RESULT	6	
ID	AARI0559	AARI0559 standard; Protein; 1962 AA.		
AG	AARI0559;			
XX				
DT	15-Apr-1991	(first entry)		
XX				
DE	Mutant protease (A137G/K138L/T139A).			
XX				
KW	Mutant protease gene; fermentation; foodstuff; flavouring;			
KM	Lactic acid bacteria.			
OS	Lactococcus lactis SK11.			
FH	Key	Location/Qualifiers		
FT	Peptide	1..187		
FT	Protein	/label= signal_peptide 188..1962		
FT		/label= mature_protein		
XX	EP411715-A.			
PM				
XX	06-FEB-1991.			
PD				
XX	02-AUG-1990;	90EP-0202113.		
XX	04-AUG-1989;	89NL-0002010.		
XX	(NEZU-) NED INST ZUIVELONDE.			
PA	Vos PAJ, Slezzen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ,			
PI	WPJ; 1991-038622/06.			
DR	N-PSDB; AAO10413.			
XX				
PT	Mutant protease gene(s) and protease(s) - derived from type I and			
PT	III protease genes from Lactococcocal strains, used in fermentation			
PT	foodstuffs and flavourings			
XX				
PS	Disclosure: Fig 1(1-7)+5(b); 29 pp; English.			
XX				
CC	The mutant protease having new cleavage specificities is			
CC	obtained by replacing three amino acids.			
CC	This mutant may then be used to prepare hybrid proteases,			
CC	the fusion being between a type I and a type III protease of			
CC	L.lactis Wg2 and SK11.			
CC	The product has modified properties, e.g. thermostability,			
CC	alkaline/acid pH stability, oxidative stability, autoprocolysis etc.			
CC	compared to the parent protease(s). The proteases can be used for			
CC	preparing products (butter cheese, human and animal foodstuffs)			

Query Match	13.0%	Score 447.5	DB 12	Length 1962
Best Local Similarity	27.3%	Pred. No. 6.3e-18		
Matches 166	Conservative 92	Mismatches 236	Indels 111	Gaps 25
OY	78	NSNKGSDYTDSEFVNKNTENPKKEKRV--YIAEFKDEKSEKAIKELSLKNTKVLVT	134	
DB	84	nkvgqgdlyvdvlyvmsaa-pasenglllrtdysstcaelqgcnkrylaaqaasvkaaveqt	142	
OY	135	-----YDRIFNGSALETTTPDNLDKIKQIEGSISSVERAKQVOPMNNHARKEIGYEEAID	187	
DB	143	qqlagesygyvvwfgstkrvrvldpklkqlagvktvrlakvyvptdakansmanvqavws	202	
OY	188	YLSKSNAPFGKRFDRGNWISNIDGTGYRHNAMTIDDAKSMFKKEDL-KGTD--K	243	
DB	203	nyk-----ykggegtvsvidsagldprthdmcrlsd--kdvrlksadvkftclvvh	251	
OY	244	NYWLSDKIPHAENY-NGKTIKEYDDORDYFDDPHGMHIAIGLGNDEPDIDIKNPNIGD	302	
DB	252	gryfsrskpyvfyngadndltidckvd-----eqhgmhvglllgangtqgd--paksyv	303	
OY	303	GIAPNAQIFSYKMYSDAG-SGFGDETFMHAIEDSIKHNVDVSVSSGP-TGTGLVGEKY	360	
DB	304	gvapeaqllamkvfnstsdtslgsatlvsaalsedaklgadvlnmslgsngnqltedp	363	
OY	361	WQATIALRKAGIPMWVAATGNVATSSSSSMDLVANNHLKMTGTGNVTRIAHEDAIAVAS	420	
DB	364	laavqananesftaavlasngstgsat--egvnrkdygylqdnemvgsptargaltvas	421	
OY	421	AKNQVFEEDKVI-----GGESEFKY--RNIGAFEDRSKITTTNEDGKAPSKLKFVY	470	
DB	422	aeentlvlgavtlvtgltqlgipetqljshndftgsfdqkkylyvkaasg-----nl	473	
OY	471	GKGQDDILIGDLRGKIAVMDR-IYTKDLKNAFKRAMDGARAINVNTVNYNRDNMTE	529	
DB	474	skgladlyla-dakgkialvkrgefsfdkqkyaga--egaaglllvnt-----dgta	523	
OY	530	LPAMGYEADEGKSOVFISGDDCYKL-WNNINPDKTEVYRNKKEDPKDLQDTPID	587	
DB	534	lpmstsltl-tlftfclgtsvgtklydvahelpoddsigvk-----	563	
OY	588	MESFNSNKPWDEKEDKRFAPDDMKELYKEDIIVPAGSTMGSPRIDLLKPDVSPAPK	647	
DB	564	-----lflamjpn--qkytedkm--sdftsgpavnsflfpdlclapgg	602	
OY	648	NIKSTLN 654		
DB	603	nltwstgn 609		
RESULT	7			
AA010557				
ID	AA010557	standard; Protein; 1962 AA.		
XX	AA010557			
AC	AA010557			
XX	15-APR-1991	(first entry)		
DT	15-APR-1991	(first entry)		
XX				
DE		Mutant protease (A137G/K138D).		
XX				
KW		Mutant protease gene; fermentation; foodstuff; flavouring;		
XX		lactic acid bacteria.		
OS		Lactococcus lactis SK11.		
XX				
EH	Key	Location/Qualifiers		
FT	Peptide	1..187		
FT		/label=signal_peptide		

FT	Protein	188..1962
FT	/label=	mature_protein
XX		
PM	EP411715-A.	
XX		
PD	06-FEB-1991.	
XX		
PE	02-AUG-1990;	90EP-0202113.
XX		
PR	04-AUG-1989;	89NL-0002010.
XX		
PA	(NEZU-) NED INST ZUIVELONDE.	
PI	Vos P.J., Slezten R.J., De Vos W.M., Kok J., Venema G., Haandrikman A.J.;	
XX		
DR	WPI; 1991-038622/06.	
DR	N-PSDB; AAQ10411.	
XX		
PT	Mutant protease gene(s) and protease(s) - derived from type I and	
PT	III protease genes from lacticococcal strains, used in fermentation	
XX	foodstuffs and flavourings	
XX		
PS	Disclosure; Fig 1.(1-7)+5(b); 29 pp; English.	
XX		
CC	The mutant protease A137C/K138D having new cleavage specificities is	
CC	obtained by replacing two amino acids.	
CC	This mutant may then be used to prepare hybrid proteases,	
CC	the fusion being between a type I and a type III protease of	
CC	L.lactis M92 and SK11.	
CC	The product has modified properties, e.g. thermostability,	
CC	alkaline/acid pH stability, oxidative stability, autolysis etc.,	
CC	comparing products (butter cheese, human and animal foodstuffs)	
CC	prepared with the aid of lactic acid bacteria.	
CC	See also AAQ10411-17 and AAQ10870-71.	
XX		
SO	Sequence	1962 AA;

Query Match	12.8%	Score 440.5	DB 12	Length 1962
Best Local Similarity	27.2%	Pred. No. 1.6e-17		
Matches 165	Conservative 92	Mismatches 239	Indels 111	Gaps 25

```

0y 78 NSNSQCGYTSFPNKNNTENRKKDKV---YINEFDKSGEKAILESSLKNTRYLT 134
Db 84 nkvgqgdlyvdvilygmsaa-pasengillrtcdysaeiqetnkvlaagsvvaaveqt 142
-0y 135 -----YDRFNGSALETTTPRDNLIDKIKQIEGSISSVRAAKVOPRMNNHAREIGVEAID 187
Db 143 qqltagesygyvngfscfkvrvvdlprklkqglagvytvclakvyvprrtckalsmnygaws 202
0y 188 YLKSTINAFEGKNEDGRGVTSNIDTGDYRRHKAMRIDDAKASMRKKEDL-KGTD--K 243
Db 203 nyk-----yugegvvaslsgldprchkdmrlsd--kdwkllksdvekfddtlvkx 251
0y 244 NYWLSMDKTRPNAFNRY-NCKRTTYEKVDGGRDYFRPHGNMHTAGILACNDTDODLKPNFGID 302
Db 252 gryfnskpyrfynadnnltfdtkvd----eqhmhvagyligaagtgd--paksyv 303
0y 303 GIAPMAOIFYSKMYSDAG-SGFADDEIMPHAIEDSIRKHNVDSVSSEGF-TGRGLGEKY 360
Db 304 gvapeaqllamkvafnsctsgdytsatcvsaalelsaklgadvlmslgsnsgnqclde 363
0y 361 WQAIRALKKAGIPWVAVTGNVATGASSSSMDLVNHNKLKTDTGNYVTRTAANDATAIVAS 420
Db 364 laavgnaneshtaavlaagnsgtsgast--egyrnkdyjgdlnemyspsrsgattvas 421
0y 421 AKNOTVEFDKNI-----GGESFPRY--RNICAFPDKSKITTNEDGTAKPSKLFVVI 470
Db 422 aentdvilqavticldgtqlgipretllisshdfcfsdqdkfyivxdag-----nl 473
0y 471 GKGGDDDLIGLDLGKTAVMOR-IYTDOLNAFKAADDKARALMVNYNYNRDNME 529

```

```

Db 474 skgaladyta-daxghaiavrgesfiddkxyega--agaaglllvnt-----dgt a523
Qy 530 LPAMGYEADDETKSQVSISSDDGVKL--MMIMIPDKKTEYKRNNKEDFEKDLEQYYPIID 567
Db 524 tpmtsialt--tftprtglsavtqkivdwtaltpdsslqv-----5633
Qy 568 MESFNSNKNPNVGDEKEIDFKAPPTDKELKEDIIVAGSTSMGPRIDLLKPDVSAPK 647
Db 564 -----ltlamjpn---qkytedkm--sdtfsygpvsnlsfkipdltapgs 602
Qy 648 NIKSTLN 654
Db 603 niwatsqn 609

```

RESULT	8
ADP10EE0	

ID AAR10558 standard; protein; 1962 AA.

AC AAR10558;

DT 15-APR-1991 (first entry)

DE Mutant protease (A137G/K138P/T139P).

KW Mutant protease gene; fermentation; foodstuff; flavouring;

XX
XX

XX
XX
1000

FT	Peptide	1..187
FT		19601

Protein	188..1962
FT	
EM	

XX
XX
FD411715-2

XX
05-FFFF-1001

XX 03-AUG-1990: 90FEB-0303113
 DE

04-AUG-1999 09NT-0003010
XX
XX
DE

XX
XX
(WEST-) MED INST ZITVEI ONDE
DA

VOC	PAT	Sleazen	PT	No VOC	WM	Kob T	Veronema C	Haandrikman A.J.
XX								
BT								

WPB: 1991-038632/06

DR N-PSDB; AAQ10412
XY

PT Mutant protease gene(s) and protease(s) - derived from type I and TII protease genes from lacticoccal strains used in fermentation

PT foodstuffs and flavourings

PS Disclosure; Fig 1(1-1)+5(b); 29 pp; English.
 YX

The mutant protease having new cleavage specificities is obtained by renaturing three amino acids.

CC This mutant may then be used to prepare hybrid proteases,

CC The product has modified properties and thermostability.

CC compared to the parent protease(s). The proteases can be used for
CC alkaline/acid pH stability, oxidative stability, autolysis etc.

CC prepared with the aid of lactic acid bacteria.

CC See also AAQ10411-1/ and AAQ10870-1/1.

SQ Sequence 1962 AA;

Query Match	12.78;	Score 438.5;	DB 12;	Length 1962;
-------------	--------	--------------	--------	--------------

Best Local Similarity 27.2%; Pred. No. 2.1e-17;
Matches 165; Conservative 92; Mismatches 239; Indels 111; Gaps 25;

```

QY 78 NSKSGODTDSVNKNTEPRKEDKVY---YIAEFKDSGSKAIKELSLNTKVLTYT 134
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 84 nkvgqddlyvdvlgmsaa-paseengllrtcdyssteaigctnkvlaaqskaaveqvt 142
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 135 -----YDRIFNGSAIETTPDLNLDKIKQIEGSSVERAOKVOPMNNHARKEIGVEAID 187
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 143 qqtagesygyvvnvgfctkvrvvdiprkqkqagkytvtlakyvrrptdakmsmavqaws 202
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 188 YLKSINAPFGKNFDGKGMVSNIDTGTDRHKKAMRIDDAKASMRKKEDL-KGTD--K 243
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 203 nyk-----ykgegtvsvsidsgidrclchkdmrlsdd--kdvklksdvекftdlvkh 251
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 244 NWYLSDKIRPAHFNY-NGCKITVEKYDDGRDYDFDPHGMHAGLAGNDTODIKNFNGID 302
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 252 gryfnskvpygfnvadndtltdckvd-----eqgmhvaaglllgangtgdd--paksyv 303
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 303 GIAPNAQIESYKMYSDAG-SGFAGDETMFHAIEDSIKHNVDVSVSSGF-TGTGLVGEKY 360
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 304 gvaapeagllamkvfnstdtsqpsatvsaiedsaklgadvlnmslgsnsgnqtlledpe 363
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 361 WQAIRALKRAGIPMVVATGNTATSSASSWDLVANNHLKMTDTGNVTRTAHEDALAVAS 420
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 364 laavpnaesgtaavlaagsnsgtsaat--egvnykdygldnemvgspsgrgattvas 421
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 421 AKNQTEPFDDVNI-----GGESFKY--RNIGAFPDKSKITTNEGCTAPSKIKFVYI 470
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 422 aentldvltgavtltdgtqglqspetqlqshdltgsfdqkkylyvkaasg-----nl 473
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 471 GKGODDGLGLDGRKIAVMDR-IYTRDLKNAFKKADKARAIMVNTVYNNRDNMTE 529
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 474 skgaladyta-dakgkialavkrgefsfdckkyaga--agaaglllynt-----dgt 523
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 530 LPMAGTEADEGTSQVFSISGDGVKL--WNMINPDKTEVKRNKNEDEKDLQYYPID 587
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 524 tpmstalt--tftptgltsvsgklydvtahpddslgvk-----563
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 588 MESFNSKKRPVSGEKEIDFFAPDTEKLYKEDLIYVAGSTSGRPRIIDLKPDVSAK 647
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 564 -----ltlamlpn--qkytedkm--sdltsygsprvsnlfskpdltap9 602
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 648 NIKSTLN 654
      | | | | |
DB 603 nlwstqn 609
      | | | | |
      ULT 9
      ID AAR10941 standard; Protein; 1968 AA.
      AC AAR10941;
      DT 15-APR-1991 (first entry)
      DE Mutant protease (delta137-139/Ins9).
      KM Mutant protease gene; fermentation; foodstuff; flavouring;
      KM lactic acid bacteria.
      XX Lactococcus lactis SK11.
      XX OS
      XX FT Key Location/Qualifiers
      FT Peptide 1..187
      FT Protein /label= sig_peptide
      FT Protein /label= mat_protein
      XX EP411715-A.
      XX PD 06-FEB-1991.
      XX

```

```

PF 02-AUG-1990; 90BP-0202113.
XX
PR 04-AUG-1989; 89NL-0002010.
XX
PA (NEZU-) NED INST ZUIVELONDE.
XX
PI Vos PAJ, Slezén RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;
DR WPI: 1991-038622/06.
XX N-PSDB; AAQ10871.
XX
PT Mutant protease gene(s) and protease(s) - derived from type I and
PT III protease genes from Lactococcal strains, used in fermentation
PT foodstuffs and flavourings
XX
PS Disclosure; Fig 1(1-7)+5(b); 29pp; English.
XX
CC The wild-type L.lactis SK11 protease gene sequence was determined
CC by the applicant (EP-307011).
CC The mutant protease having new cleavage specificities is
CC obtained by deleting three amino acids (line bps) and inserting
CC 9 other residues.
CC This mutant may then be used to prepare hybrid proteases,
CC the fusion being between a type I and a type III protease of
CC L.lactis Wg2 and SK11.
CC The product has modified properties, e.g. thermostability,
CC alkaline/acid pH stability, oxidative stability, autolysis etc.,
CC compared to the parent proteases. The proteases can be used for
CC preparing products (butter cheese, human and animal foodstuffs)
CC prepared with the aid of lactic acid bacteria.
CC See also AAQ10411-17 and AAQ10870-71.
XX
SQ Sequence 1968 AA:

Query Match 12.6%; Score 435.5; DB 12; Length 1968;
Best Local Similarity 27.1%; Pred. No. 3.2e-17;
Matches 166; Conservative 92; Mismatches 238; Indels 117; Gaps 26;

QY 78 NSKSGODTDSVNKNTEPRKEDKVY---YIAEFKDSGSKAIKELSLNTKVLTYT 134
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 84 nkvgqddlyvdvlgmsaa-paseengllrtcdyssteaigctnkvlaaqskaaveqvt 142
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 135 -----YDRIFNGSAIETTPDLNLDKIKQIEGSSVERAOKVOPMNNHARKEIGVEAID 187
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 143 qqtagesygyvvnvgfctkvrvvdiprkqkqagkytvtlakyvrrptdakmsmavqaws 202
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 188 YLKSINAPFGKNFDGKGMVSNIDTGTDRHKKAMRIDDAKASMRKKEDL-KGTD--K 243
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 203 nyk-----ykgegtvsvsidsgidrclchkdmrlsdd--kdvklksdvекftdlvkh 251
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 244 NWYLSDKIRPAHFNY-NGCKITVEKYDDGRDYDFDPHGMHAGLAGNDTODIKNFNGID 302
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 252 gryfnskvpygfnvadndtltdckvd-----eqgmhvaaglllgangtgdd--paksyv 303
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 303 GIAPNAQIESYKMYSD--AG-SGFAGDETMFHAIEDSIKHNVDVSVSSGF-TGTG 354
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 304 gvaapeagllamkvfnstdtsqpsatvsaiedsaklgadvlnmslgsnsgnqtlledpe 363
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 355 LGEKYYQAIRALKRAGIPMVVATGNTATSSASSWDLVANNHLKMTDTGNVTRTAHED 414
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 364 tledpelaavpnaesgtaavlaagsnsgtsaat--egvnykdygldnemvgspsgrg 421
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 415 AIAVASKNTEPFDDVNI-----GGESFKY--RNIGAFPDKSKITTNEGCTAPSK 464
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 422 atvvasenldvltgavtltdgtqglqspetqlqshdltgsfdqkkylyvkaasg---- 477
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 465 LKEFYVIGKODDGLGLDGRKIAVMDR-IYTRDLKNAFKKADKARAIMVNTVYNN 523
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 478 -----nlsgaladyta-dakgkialavkrgefsfdckkyaga--agaaglllynt----- 525
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 524 RDMWTELPAMGYEADGTSQVFSISGDGVKL--WNMINPDKTEVKRNKNEDEKDL 581
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Db 536 --dgtatpmtslalt--ttfptfglssvlgqklvdwvtahpddslyvk----- 569
 Oy 582 QYPTDMSPFNSNKPNGVDEKEIDKFAPDIDKELYKEDIIVPAGSTSMGPRIDLLKPD 641
 Db 570 -----lclamipn---qkytedkm--sdftsygypvsnlsfkp 602
 Oy 642 VSAPGKNIKSTLN 654
 Db 603 ltapggnlwtstgn 615

RESULT 10

AA010562
 ID AAR10562 standard; Protein: 1959 AA.

AC AAR10562;

DT 15-APR-1991 (first entry)

DE Mutant protease (delta137-139).

FM Mutant protease gene; fermentation; foodstuff; flavouring;
 lactic acid bacteria.

OS Lactococcus lactis SK11.

Key Location/Qualifiers

FT Peptide

FT /label= signal_peptide

FT Protein

FT /label= mature_peptide

PN EP411715-A.

PD 06-FEB-1991.

PF 02-AUG-1990; 90EP-0202113.

PR 04-AUG-1989; 89NL-0002010.

PA (NEZU-) NED INST ZUIVELONDE.

PI Vos PAJ, Slezzen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;

DR WPI: 1991-038622/06.

DR N-PSDB; AAQ10416.

PS Mutant protease gene(s) and protease(s) - derived from type I and

FT III protease genes from lactococcal strains, used in fermentation

FT foodstuffs and flavourings

XX Disclousure; Fig 1(1-7)+5(b); 29 pp; English.

XX The mutant protease having new cleavage specificities is

CC obtained by deleting three amino acids.

CC This mutant may then be used to prepare hybrid proteases,

CC the fusion being between a type I and a type III protease of

CC L. lactis Wg2 and SK11.

CC The product has modified properties, e.g. thermostability,

CC alkaline/acid pH stability, oxidative stability, autolysis etc.,

CC compared to the parent protease(s). The proteases can be used for

CC preparing products (butter cheese, human and animal foodstuffs)

CC prepared with the aid of lactic acid bacteria.

CC See also AAQ10411-17 and AAQ10870-71.

XX Sequence 1959 AA;

Query Match 12.6%; Score 435; DB 12; Length 1959;

Best Local Similarity 27.1%; Pred. No. 3.4e-17;

Matches 164; Conservative 92; Mismatches 238; Indels 112; Gaps 25;

78 NSNKSQGDYTSFVNKNTENPKEDKV--YIAEFKDESGEKAIKELSLKNTKVLVT 134

Db 84 nkvqgdilvdvliqmsaa--pasengllrtdysslaeqgetlnkvlaaqasvkaaveqvt 142
 Oy 135 -----YDRIFNGSAITETPDNDLKIKQIEGISSVEAQAQVPMNINAREIEVEEID 187
 Db 143 qtagesygyvvnvfstvrvvdiplkqlqigvntvltakkyuplclakansmanvqaws 202
 Oy 188 YLKSINAFGKNFPGRCGMVINISIDGTDRHKAHRIIDDAKASRFFKEDL-KGTD--K 243
 Db 203 nyk-----ykgegtlvsvidsqldphtkmltsdd--kdvkltsdvtektfdvkh 251
 Oy 244 NYWLSDKIPHAENY--NGCKITVEKYDDGRDYFDPHGHIAGIAGNDEODIKNFNGID 302
 Db 252 gryfnskvpygfnyadndtlddkvd-----eqhgvhvgilgangtgd--paksyv 303
 Oy 303 GIAPRAQIFSKYKMSDAGSGFAGDETMFAIEDSIKHNVDVVSSSG--TGTGLVGEKYM 361
 Db 304 gvapeaqlamkvfns--sdtsaaltvsaiedsaklgadvlmslgansgnqtlep 361
 Oy 362 QAIRALKRAGIPNVVATGNTVATSSSSMDLVANNHKKMTDTGNVTRTAHEDAIAVASA 421
 Db 362 aavgnanesgtaavlsagngsgtsat--egvnykdyglqdnemvsgpgrgaltvasa 419
 Oy 422 KNOITVEFDKNI-----GSESFKY--RNIGAFEDKSKITTNEDGTAKSLKPYIIG 471
 Db 420 entdvlqavltldgtlqldpeltqlshdftgsfdqkfyivkdaag-----nls 471
 Oy 472 KGPQODLIGDLRKRIANMDR-ITTKDLKNAFKKAMDGAIAIVNVNINRDNMTEL 530
 Db 472 kgaladyta-dakgialvktgeisfdckqyaga--agaagllivnt-----dgtat 521
 Oy 531 PAMGYEADGCTKSOFSISGDDGVKL--WMNINPDKTEVRNNKEDFKLEQYYPIDM 588
 Db 522 pmtslalt--ttfptfglssvlgqklvdwvtahpddslyvk----- 560
 Oy 589 ESFNSNKPNGVDEKEIDKFAPDIDKELYKEDIIVPAGSTSMGPRIDLLKPDVSAPGN 648
 Db 561 -----lclamipn---qkytedkm--sdftsygypvsnlsfkpdltapgn 600
 Oy 649 IKSTLN 654
 Db 601 lwtstgn 606

RESULT 11

AA010560
 ID AAR10560 standard; Protein: 1962 AA.

AC AAR10560;

DT 15-APR-1991 (first entry)

DE Mutant protease (K138N).

FM Mutant protease gene; fermentation; foodstuff; flavouring;
 lactic acid bacteria.

OS Lactococcus lactis SK11.

Key Location/Qualifiers

FT Peptide

FT /label= signal_peptide

FT Protein

FT /label= mature_protein

PN EP411715-A.

PD 06-FEB-1991.

PF 02-AUG-1990; 90EP-0202113.

PR 04-AUG-1989; 89NL-0002010.


```

QY 188 YLKSINAFPGKNDGKGVNISIDTGTDRHKAHRIDDDAKASRFKKEDL-KGTD---K 243
D 203 nyk-----ykgeglvsvldsgldprtkhmlsdd--kdklksdvckftdckh 251
QY 244 NYWLSDRIPAHFNY--NGKLTVEKYDDGRDYPFHGHMHLIGLIGAGNTEODIKNFNGID 302
D 252 gryfnskvpygfyndndlttdckvd-----eqhgmhvagiigangtgdd--paksyv 303
QY 303 GIAPNAOIFSKMYSDAG--SGFADGEMFHAIEDSIKHNVDVSVSSGF-TGTGLVGEKY 360
D 304 gvapeaqlamkvfsnctsanctsaclvsaiedsaklgadvlmslgsdsngtledpe 363
QY 361 WQAIRALKRAGIPMVVATGNVATSSASSMDLVANNHLMKMTDGNVTRTAHEDAIIVAS 420
D 364 laavqanessglaaviasgngtsat--egvnhkdygylqdnemvgsprgtsatlvass 421
QY 421 AKNOTVEFDKVN-----GGESFNY--RNIGAFDPSKITTNEGDKAPSKLFEVYI 470
D 422 aentdvltqavltldgtglqigpeltqlshtfcsfgdkfyivkdasg-----nl 473
QY 471 GKGDODLIGDLRGKIAVMDR-IYTKDLKNAFKKAMDKGARAIWVNTVYNRDMNTE 529
D 474 skgaladyta-dakgkialavtrgefsfdckkkyaga--agaagllivnt-----dgt 523
QY 530 LPAMGYEADGEGTSQVFSISGDDGVKL--WNMNPDKKTEVKRNKKEDKDLQYYPID 587
D 524 tpmtsialt--ctfptfglsvtgkivdwvtahpddslgyk----- 563
QY 588 MESFNSNKPNGDEKEIDFKRAPDTEKELYKEDIIVPAGTSWGPRIIDLKPPVSAAGK 647
D 564 -----ltlamlpn--qkyledkm--sdfcsygvpsnlstfkpdlctapgg 602
QY 648 NIKSTLN 654
D 603 niwstgn 609

RESULT 13
AA10563
ID AA10563 standard; Protein; 1962 AA.
AC AA10563;
XX
XX 15-APR-1991 (first entry)
XX
XX Mutant protease (K748T).
XX
XX Mutant protease gene; fermentation; foodstuff; flavouring;
XX lactic acid bacteria.
XX Lactococcus lactis SK11.
XX
XX Key Location/Qualifiers
XX Peptide 1..187
XX FT /label= signal_peptide
XX FT /label= 188..1962
XX FT /label= mature_protein
XX
XX EP411715-A.
XX
XX 06-FEB-1991.
XX
XX 02-AUG-1990; 90EP-0202113.
XX
XX 04-AUG-1989; 89NL-0002010.
XX
XX (NEZU-) NED INST ZUIVELONDE.
XX
XX Vos PAJ, Slezzen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;
XX
XX WPI: 1991-038622/06.
XX
XX N-PSDB: AA010417.
XX

```

```

PT Mutant protease gene(s) and protease(s) - derived from type I and
PT II protease genes from lactococcal strains, used in fermentation
PT foodstuffs and flavourings
XX
XX PS Disclosure; Fig 1(1-7)+5(a); 29 pp; English.
XX
XX CC The mutant protease K748T having new cleavage specificities is
XX obtained by carrying out single amino acid substitutions.
XX CC This mutant may then be used to prepare hybrid proteases,
XX CC the fusion being between a type I and a type II protease of
XX CC L.lactis Wg2 and SK11.
XX CC The product has modified properties, e.g. thermostability,
XX CC alkaline/acid pH stability, oxidative stability, autolysis etc.,
XX CC compared to the parent protease(s). The proteases can be used for
XX CC preparing products (butter cheese, human and animal foodstuffs)
XX CC prepared with the aid of lactic acid bacteria.
XX
XX CC See also AA010411-17 and AA010870-71.
XX
XX
XX SQ Sequence 1962 AA:

```

```

Query Match 12.68; Score 434.5; DB 12; Length 1962;
Best Local Similarity 27.08; Pred. No. 3.0e-17;
Matches 164; Conservative 92; Mismatches 240; Indels 111; Gaps 25;

```

```

QY 78 NSNKSQGDYDSEFYNNKTENPKEDKVY---YIAEFKDESGEKAIEKLSLAKTKVLYT 134
D 84 nkvgqgdlyvndvlyqmsaa--pasengllrtidsytsaetqetlnvlaqasvkaaveqvt 142
QY 135 -----YDRIFNGSAIETTPDNLKIKOLEIGISVERBAQVOPMMNHARKEIVEAID 187
D 143 qtagesygyvvnvgsfckvrvndiprkikqlgavlvtlakvypcldakansmanvqavws 202
QY 188 YLKSINAFPGKNDGKGVNISIDTGTDRHKAHRIDDDAKASRFKKEDL-KGTD---K 243
D 203 nyk-----ykgeglvsvldsgldprtkhmlsdd--kdklksdvckftdckh 251
QY 244 NYWLSDRIPAHFNY--NGKLTVEKYDDGRDYPFHGHMHLIGLIGAGNTEODIKNFNGID 302
D 252 gryfnskvpygfyndndlttdckvd-----eqhgmhvagiigangtgdd--paksyv 303
QY 303 GIAPNAOIFSKMYSDAG--SGFADGEMFHAIEDSIKHNVDVSVSSGF-TGTGLVGEKY 360
D 304 gvapeaqlamkvfsnctsanctsaclvsaiedsaklgadvlmslgsdsngtledpe 363
QY 361 WQAIRALKRAGIPMVVATGNVATSSASSMDLVANNHLMKMTDGNVTRTAHEDAIIVAS 420
D 364 laavqanessglaaviasgngtsat--egvnhkdygylqdnemvgsprgtsatlvass 421
QY 421 AKNOTVEFDKVN-----GGESFNY--RNIGAFDPSKITTNEGDKAPSKLFEVYI 470
D 422 aentdvltqavltldgtglqigpeltqlshtfcsfgdkfyivkdasg-----nl 473
QY 471 GKGDODLIGDLRGKIAVMDR-IYTKDLKNAFKKAMDKGARAIWVNTVYNRDMNTE 529
D 474 skgaladyta-dakgkialavtrgefsfdckkkyaga--agaagllivnt-----dgt 523
QY 530 LPAMGYEADGEGTSQVFSISGDDGVKL--WNMNPDKKTEVKRNKKEDKDLQYYPID 587
D 524 tpmtsialt--ctfptfglsvtgkivdwvtahpddslgyk----- 563
QY 588 MESFNSNKPNGDEKEIDFKRAPDTEKELYKEDIIVPAGTSWGPRIIDLKPPVSAAGK 647
D 564 -----ltlamlpn--qkyledkm--sdfcsygvpsnlstfkpdlctapgg 602
QY 648 NIKSTLN 654
D 603 niwstgn 609

RESULT 14
AA10940
ID AA10940 standard; Protein; 1974 AA.

```

```

XX AC AAR10940;
XX 15-APR-1991 (first entry)
XX DE Mutant protease (delta137-139/Ins15).
XX KM Mutant protease gene; fermentation: foodstuff; flavouring;
XX KM lactic acid bacteria.
XX OS Lactococcus lactis SK11.
XX FH Key Location/Qualifiers
XX FT Peptide 1..187
XX FT /Label- SIG_PEPTIDE
XX FT Protein 188..1974
XX FT /Label- MAT_PROTEIN
XX EP411715-A.
XX 06-FEB-1991.
XX PE 02-AUG-1990; 90EP-0202113.
XX PR 04-AUG-1989; 89NL-0002010.
XX PA (NEZU-) NED INST ZUIVELONDE.
XX PI Vos PAJ, Stezen RJ, De Vos WM, Kok J, Venema G, Haandrikman AJ;
XX DR WPI; 1991-038622/06.
XX DR N-PSDB; AAO10870.
XX PT Mutant protease gene(s) and protease(s) - derived from type I and
XX PT III protease genes from lactococcal strains, used in fermentation
XX PT foodstuffs and flavourings
XX PS Disclosure: Fig 1(1-7)+5(b); 29pp; English.
XX CC The wild-type L. lactis SK11 protease gene sequence was determined
XX CC by the applicant (EP-307011).
XX CC The mutant protease having new cleavage specificities is
XX CC obtained by deleting three amino acids (nine bps) and inserting
XX CC 15 other residues.
XX CC This mutant may then be used to prepare hybrid proteases,
XX CC the fusion being between a type I and a type III protease of
XX CC L. lactis Wg2 and SK11.
XX CC The product has modified properties, e.g. thermostability,
XX CC alkaline/acid pH stability, oxidative stability, autolysis etc.,
XX CC compared to the parent protease(s). The proteases can be used for
XX CC preparing products (butter cheese, human and animal foodstuffs)
XX CC prepared with the aid of lactic acid bacteria.
XX CC See also AAO10411-17 and AAO10870-71.
XX CC
XX SQ Sequence 1974 AA;

```

```

Query Match 12.5% Score 431.5; DB 12; Length 1974;
Best Local Similarity 26.8%; Pred. NO. 5.5e-17;
Matches 166; Conservative 92; Mismatches 238; Indels 123; Gaps 26;

```

```

OY 78 NSNKSOGDYTFVNKNTKPKKEDKV--YIAEFKDESEKAIKELSLKNTKVLXT 134
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 84 nkvgqddlyvdylvqmsaa-pasengllrtfdysstaetqetnkvlaaasvkaavegt 142
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 135 -----YDRIRGSAIETTPDNLDKIOEGISSVERAKQOPMNMHAKKEIGVEAID 187
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 143 qqtagesygyvngfctkryvvdiprlkqlagvktvclakvyprtdakansmanyavva 202
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 188 YLKSINAPGKNPDGRCGVYISNIDGTDRHKAMRIDDAKSMRPKEDL-KGTP---K 243
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 203 nyk-----ykgegtlvsvldsgldpchkdmrlsdd--kdvkltsdvekfctlvkh 251
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

OY 244 NWLSDKIPHAENFY-NGSKITVEKYDGRDYFDPHGHIACILAGNTEQDIKPNKCID 302
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 252 gryfnskypgyfngadndncltdkvd-----eqhgmhvagllgangegdd--paksav 303
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 303 GIAPNAOIFSYKMSD-----AG-----SGFAGDETFMFAIEDSIKHNVDVYSSSG 349
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 304 gvapeeqllamvrfnsdtsqktglagktgkcgkgsaalvsaileosaikgadvlnmslg 363
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 350 F-TGTGLVEKRYWQAIRALKAGIPWVATGVNATSSASSMDLVANHLKMTDTGMVTR 408
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 364 snsgnqtleddpelaavqnaesgtaavisaagngtsast--egynkdyjgldnemys 421
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 409 TAAHEDALIVASAKQTVBFEDKVN-----GGESEFY--RMICAFEDSKITTTNEDG 458
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 422 pgtsgattvsaadentdvltgavltgltgltqipreqltqshdfgsgdkkfyivkda 481
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 459 TKAPSKLKFVYIGKGDOLLIGDLRGKIAVWDR-IYTTDLKNAFKKADKARAIMVYN 517
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 482 sg-----nlskgaladyta-dakgkialvkrgefsiddkkyaqda--agaaglllvn 530
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 518 TVNYNRDNWTELPAWGYBDEGTSQVPSISGDDGVKL--NMIMPDKTEVKRNKED 575
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 531 t-----dgtatpmsialt--tlfpctglssvgqklvdvltahpddslyvk----- 575
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 576 FKDKLEQYYPIDMESFNSKNKPNVGDEKEIDFKFAPDTOKELYKEDIIVAGSTSMGPRID 635
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 576 -----ltlamlpn---qkytedkm--sdfstsygpran 602
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 636 LLLKPDVSAPEGKNIKSTLN 654
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 603 lsfkpdltapgsnltwstgn 621
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

```

RESULT 15
AAP94145
ID AAP94145 standard; protein; 1560 AA.
XX
XX AAP94145;
XX AC AAP94145;
XX XX
XX DT 07-JUN-1990. (first entry)
XX DE S. cremoris protease.
XX KM Lactic acid bacteria; cheese; Streptococcus cremoris SK112;
XX KM protease; pK112; chymosin; prochymosin.
XX PN NL8701378-A.
XX XX
XX PD 02-JAN-1989.
XX XX
XX PF 12-JUN-1987; 87NL-0001378.
XX XX
XX PR 12-JUN-1987; 87NL-0001378.
XX XX
XX PA (NEZU-) NEDERL INS ZUIVELON.
XX PI
XX PI Simons AFM, De Vos WM;
XX DR WPI; 1989-030097/04.
XX DR N-PSDB; AAN91159.
XX PT DNA fragment having region specific for lactic acid bacteria -
XX PT is contained in plasmid in microorganism used in prodn. of
XX PT protein and food prodn. eg cheese.
XX PS Claim 6; fig 7a; 43pp; Dutch.
XX XX
XX CC The DNA encoding prochymosin can be cloned into a plasmid (esp from
XX CC S. cremoris SK112) and used to produce large amounts of the protein by
XX CC recombinant DNA techniques. The protein synthesis is driven by the
XX CC regulatory region of this protease. This could overcome the shortage
XX CC of prochymosin due to a shortage of calf stomachs and increasing cheese
XX CC prodn. Prochymosin is also used in prodn. of yoghurt, butter and

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: December 7, 2001, 11:45:25 ; Search time 14.18 Seconds
(without alignments)
1053.751 Million cell updates/sec

Title: US-09-590-991-6
Perfect score: 3454
Sequence: 1 PVPVLAADTSSEDAALNISDK.....PGNNIKSTLVNIGKSTGY 664

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues
Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfillsl.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3269	94.6	1007	4 US-08-961-083-216	Sequence 216, App
2	302	8.7	1164	2 US-08-589-756-1	Sequence 1, App11
3	302	8.7	1164	4 US-09-206-800-1	Sequence 1, App11
4	296.5	8.6	1150	2 US-08-589-756-3	Sequence 3, App11
5	296.5	8.6	1150	4 US-09-206-800-3	Sequence 3, App11
6	293	8.5	1167	2 US-08-589-756-2	Sequence 2, App11
7	293	8.5	1167	4 US-09-206-800-2	Sequence 2, App11
8	273	7.9	383	4 US-09-206-800-8	Sequence 8, App11
9	273	7.9	383	4 US-09-206-800-11	Sequence 8, App11
10	268	7.8	383	4 US-09-206-800-6	Sequence 11, App1
11	268	7.8	383	4 US-09-206-800-9	Sequence 6, App11
12	264	7.6	383	4 US-09-206-800-7	Sequence 9, App11
13	264	7.6	383	4 US-09-206-800-10	Sequence 10, App1
14	236	6.8	654	4 US-08-894-818B-35	Sequence 35, App1
15	229.5	6.6	659	4 US-08-894-818B-1	Sequence 1, App11
16	214.5	6.2	659	4 US-08-894-818B-5	Sequence 5, App11
17	201	5.8	522	4 US-08-894-818B-3	Sequence 3, App11
18	196	5.7	382	2 US-08-504-265B-75	Sequence 75, App1
19	181.5	5.3	381	1 US-07-772-087-5	Sequence 5, App11
20	181.5	5.3	381	1 US-08-322-965-2	Sequence 2, App11
21	180	5.2	382	1 US-08-460-343B-2	Sequence 2, App11
22	180	5.2	382	1 US-08-398-028B-2	Sequence 2, App11
23	180	5.2	382	1 US-08-504-265B-2	Sequence 1, App11
24	180	5.2	903	1 US-08-750-532-1	Sequence 1, App11
25	180	5.2	1398	1 US-08-750-532-9	Sequence 9, App11
26	180	5.2	1398	1 US-08-894-818B-8	Sequence 8, App11
27	177	5.1	352	1 US-07-923-260A-1	Sequence 1, App11

28	177	5.1	352	1 US-07-923-260A-6	Sequence 6, App11
29	177	5.1	379	1 US-08-685-774-2	Sequence 2, App11
30	175.5	5.1	378	1 US-07-772-087-4	Sequence 4, App11
31	175	5.1	379	2 US-08-845-295A-1	Sequence 1, App11
32	175	5.1	379	3 US-09-140-933-1	Sequence 1, App11
33	175	5.1	379	4 US-09-146-661-1	Sequence 1, App11
34	175	5.1	379	4 US-09-150-515-1	Sequence 1, App11
35	174	5.0	352	1 US-07-923-260A-3	Sequence 3, App11
36	173.5	5.0	350	1 US-07-923-260A-4	Sequence 4, App11
37	173	5.0	382	4 US-09-255-502-2	Sequence 2, App11
38	173	5.0	382	6 5472855-2	Patent No. 5472855
39	171	5.0	352	1 US-07-923-260A-2	Sequence 2, App11
40	171	5.0	380	6 RE34606-2	Patent No. RE34,60
41	169.5	4.9	381	1 US-07-772-087-6	Sequence 6, App11
42	169	4.9	381	1 US-08-173-508-12	Sequence 12, App1
43	169	4.9	381	2 US-08-265-510-12	Sequence 12, App1
44	169	4.9	381	3 US-08-951-742-12	Sequence 12, App1
45	168.5	4.9	380	3 US-09-049-867-2	Sequence 2, App11

ALIGNMENTS

```
RESULT 1
US-08-961-083-216
: Sequence 216, Application US/08961083
: Patent No. 6159469
:
: GENERAL INFORMATION:
: APPLICANT: Choi et. al.
: TITLE OF INVENTION: Streptococcus pneumoniae Antigen and Vaccines
: NUMBER OF SEQUENCES: 452
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,083
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 216:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1007 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-961-083-216
```

```
Query Match 94.6% Score 3269 DB 4 Length 1007
Best local Similarity 100.0% Pred. No. 2.8e-209
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0
0Y 38 ETSDFFKKTAVTIKEVSKNPVINDNTSNEBAKIKENSNSKSGDYDTSFVNKNTEN 97
|||||
```

```

Db      1 ETSDFEKKRTAVIKEKEVSVKNPIDNNTSNEAKIKEENSNNKSGQDYSFVNKNTEN 60
QY      98 PKKEDKVVYIAEFKDKESGKAIKELSLKNTKVLTYTDRIFNCSAIEPTDNLKIQI 157
Db      61 PKKEDKVVYIAEFKDKESGKAIKELSLKNTKVLTYTDRIFNCSAIEPTDNLKIQI 120
QY      158 EGISSVERAOKVOPMMNHARKEIGVEEAIDYLSINAPFGKNFDRGVNISNIDGTDR 217
Db      121 EGISSVERAOKVOPMMNHARKEIGVEEAIDYLSINAPFGKNFDRGVNISNIDGTDR 180
QY      218 HKARIDDAKASMRFKEDLKGTDKNWLSDKIPHAFTNYNGKITYEKTDDGADYDP 277
Db      181 HKARIDDAKASMRFKEDLKGTDKNWLSDKIPHAFTNYNGKITYEKTDDGADYDP 240
QY      278 HGMIHAGIAGNDTEODIKNFNGIDGIAFNAQIFSYKMYNAGSGFAGDETFHAIEDSI 337
Db      241 HGMIHAGIAGNDTEODIKNFNGIDGIAFNAQIFSYKMYNAGSGFAGDETFHAIEDSI 300
QY      338 KHNVDVSVSSGFTGTGLVGEKYQWAIKALRAGIIPVYATGNVATSSSSMDLVANNH 397
Db      301 KHNVDVSVSSGFTGTGLVGEKYQWAIKALRAGIIPVYATGNVATSSSSMDLVANNH 360
QY      398 LKMTDGTAVTAAHEDAIANAASAKNQTVEFDKNVIGESFRYRNIGAFDOKSKITTED 457
Db      361 LKMTDGTAVTAAHEDAIANAASAKNQTVEFDKNVIGESFRYRNIGAFDOKSKITTED 420
QY      458 GTRAPSKLKFYIIGKQODLIGDLRGKIAVMDRITTKDLNAFKKAMDGARIMVYN 517
Db      421 GTRAPSKLKFYIIGKQODLIGDLRGKIAVMDRITTKDLNAFKKAMDGARIMVYN 480
QY      518 TVNYYNRDNTLTPAMGEAEDEGTRKSOVFSISGDGVKLMNINPDKKTEYVRNNKEDFK 577
Db      481 TVNYYNRDNTLTPAMGEAEDEGTRKSOVFSISGDGVKLMNINPDKKTEYVRNNKEDFK 540
QY      578 DKEQYVYIDMESNSKPNVNGDEKEIDFKAPDTEKLYKEDITYVPGSTSWGRIDL 637
Db      541 DKEQYVYIDMESNSKPNVNGDEKEIDFKAPDTEKLYKEDITYVPGSTSWGRIDL 600
QY      638 LKPDVSAPGKNIKSTLVYNGKSTYGY 664
Db      601 LKPDVSAPGKNIKSTLVYNGKSTYGY 627

RESULT 2
: Sequence 1, Application US/08589756
: Patent No. 5846547
: GENERAL INFORMATION:
: APPLICANT:
: TITLE OF INVENTION: STREPTOCOCCAL C5A PEPTIDASE VACCINE
: NUMBER OF SEQUENCES: 5
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/589,756
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1164 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-589-756-1

```

```

Query Match      8.7%; Score 302; DB 2; Length 1164;
Best Local Similarity 24.3%; Pred. No. 4.1e-12;
Matches 146; Conservative 90; Mismatches 194; Indels 172; Gaps 32;
QY      97 NPKKEDKVVYIAEFKDKESGKAI-----KELSLKNTKVLTYTDRIFNCSAIEPTT 147

```

```

Db      24 NAQSDIKANTVTE--DTPATEQAVETTPOTTVSEEVPSSEKTKTPQTPD-----DAEETV 76
QY      148 PDNLDKTIQIGISSEVEAOKVOPMMNHARKEIGVEEAID--YLSKINAPFGKNFDGGM 205
Db      77 ADDANDL-----APQAPAKTDPST--ATSKATIRLDNDSQVYTLQEKAGK--GAGT 124
QY      206 VISNIDGTDRHKKARIDDDAKASMRFKEDLKGTDKN-----WLSDKIPHAFTNYN 259
Db      125 VVAVIDAGFDKNHHEMRITDKAKARYQ--SKEDLEAKKHEHGTIGGEWVNDVAYHYDHSK 183
QY      260 GGTATVEKYDGRDYEDPHGHIAGIAGNDTEODIKNFNGIDGIAFNAQIFSYKMYN----- 315
Db      184 DGKTAVDQ-----HGTNHSGLISGN--APSETEKPYRLEGAMPQALLMLNVEIYN 233
QY      316 -YSDAGSGFAGDETFHAIEDSIKHNVDVSVSSG---FTGTGLVGEKYQWAIKALRAGK 371
Db      234 GLADYARNYA-----QAIRDAVNLGARVYNNISFQNALAYANLPDETKKPPVAKSK--G 286
QY      372 IPVYATGNVATSSSSMDLVANNHLMKMTDGTAVTAAHEDAIANA--SAKNQTVFED 429
Db      287 VRIYTTAGNDSSFGKTRPLA--DH---PDYGVGTFAAADSTLTVAISGPDNLTEETA 341
QY      430 KVNIGESSEFKYRNIGAFDOKSKITTEDGTRAPSK-LKFVYIGKQODLIGDLRGKIA 488
Db      342 MYKTDQDQDK-----EMFVLSTNR--FEPKNAYDYAVANRGMKEDDF-KDYKGRKIA 389
QY      489 VMDRIT--YTNDLNAFKKAMDGARIMVYNTVYVNRD-----NMTLTPAMGEA 537
Db      390 LIERSDIDFTDKIANA--KKA---GAVGVLI---YDQDQGFPIELPNVQMPAA----- 436
QY      538 DEGTRKSOVFSISGDGVKLMNINPDKKTEYVRNNKEDFKLEQYVYIDMESFNSKPN 597
Db      437 -----FISRKDGLL-----KD----- 448
QY      598 VGDEKEIDFKAPDTEKLYKEDITYVPGS-----TSWGRIDLKLPVSAFGKNIK 650
Db      449 -NSQKTIFFNATPK-----VLPTASGTKLSRFSSWGLTADGNIKPIAAGQDIL 497
QY      651 ST 652
Db      498 SS 499

RESULT 3
: Sequence 1, Application US/09206800
: Patent No. 6270775
: GENERAL INFORMATION:
: APPLICANT: Cleary, P. P.
: TITLE OF INVENTION: Streptococcal C5a peptidase vaccine
: FILE REFERENCE: 600.349US2
: CURRENT APPLICATION NUMBER: US/09/206,800
: CURRENT FILING DATE: 1998-12-07
: EARLIER APPLICATION NUMBER: US 08/589,756
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 1
: LENGTH: 1164
: TYPE: PRT
: ORGANISM: Streptococcus pyogenes, strain 49
: US-09-206-800-1

```

```

Query Match      8.7%; Score 302; DB 4; Length 1164;
Best Local Similarity 24.3%; Pred. No. 4.1e-12;
Matches 146; Conservative 90; Mismatches 194; Indels 172; Gaps 32;
QY      97 NPKKEDKVVYIAEFKDKESGKAI-----KELSLKNTKVLTYTDRIFNCSAIEPTT 147
Db      24 NAQSDIKANTVTE--DTPATEQAVETTPOTTVSEEVPSSEKTKTPQTPD-----DAEETV 76

```

QY	148	PNDLCKIKQIEGISSVEYERAKQVQPMNNAKREIGVEEAD--YLSINAPFEKNGDEGGM	205
Db	77	ADDANDL-----APAPAKPTPPTS-ATSAATIRLDLNDPSQVATLOEKAGK---GAGT	124
QY	206	VISNIDTGTDRHKAMRIDDDAKASMRKKEDLKGTDKNY-----WLSDKIPAFNYYN	259
Db	125	VVAVIDAGFDKKNHEAMRLTDKAKARYQ-SKEDLEKAKKEHGITTGEMVNDKVAAYHYDSK	183
QY	260	GKKTVEKYYDDGRDYFDPHGHMILAGLTNGTEQDIKNFNGIDGIAFPNAQITSYKM----	315
Db	184	DGKTAIVDD-----BHGTHVSGILSGN-APSEKPEYRLLEGAMPQAQLLMREYIVN	233
QY	316	-YSDAGSEFAGDETFMFAIEDIRKHNVDVSVSSG---FTGTGLVGEXYQAIRALRKAG	371
Db	234	GLADYARNYA-----QAIRDVAINGAKYINNPFMAALATVNLNDETFKKPPVYAKSK-G	286
QY	372	IPMVVATGNVATSASSSSWMDLVANNHLMKMTDGTGNTRTAAHEDAIYA--SAKNOTVEED	429
Db	287	VRIYVTAGNDSSFGKTRPLA--DH---PDYGVGTPAADSTLTVASYSDDNQLTEA	341
QY	430	KVNIGGSEFKRNIGAFEDKSKITTNEDGTAKPSK-LKFVYIGKQODDLISLDRGKIA	488
Db	342	MKTTDDQDDK-----EMPLSTNR--FEPNKAYDYAYANRMKEDDF-KDYGKZIA	389
QY	489	VMDRI---YTKDLKNAFKKAMDKGARAIIMVTVANYND-----NWTPLPMGYEA	537
Db	390	LIERSIDFTDKTIAA--KKA--GAGVLT-----YDODKGFPIELPVDWMPAA----	436
QY	538	DEGTSQVFSISIGDDGVKLMNMINPDKKTEYKRNKKEDFKDLQIYYPIDMESFNKNPN	597
Db	437	-----FISRKGGLL-----KD-----	448
QY	598	VGDEKEIDFKFAPDPTDKELYKEDIIVPAGS-----TSWGRIDLCLKPDVSAPGKNIK	650
Db	449	-NSQKTIFFNATPK-----VLPTASGTKLSRFSSWGLTAGDNIKPIDAIAGDIL	497
QY	651	STP 652	
Db	498	SS 499	
<p>RESULT 4</p> <p>US-08-589-756-3</p> <p>Sequence 3, Application US/08589756</p> <p>Patent No. 5846547</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT:</p> <p>TITLE OF INVENTION: STREPTOCOCCAL C5a PEPTIDASE VACCINE</p> <p>NUMBER OF SEQUENCES: 5</p> <p>COMPUTER READABLE FORM:</p> <p>MEDIUM TYPE: Floppy disk</p> <p>COMPUTER: IBM PC compatible</p> <p>OPERATING SYSTEM: PC-DOS/MS-DOS</p> <p>SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)</p> <p>CURRENT APPLICATION DATA:</p> <p>APPLICATION NUMBER: US/08/589, 756</p> <p>INFORMATION FOR SEQ ID NO: 3:</p> <p>SEQUENCE CHARACTERISTICS:</p> <p>LENGTH: 1150 amino acids</p> <p>TYPE: amino acid</p> <p>STRANDEDNESS: single</p> <p>TOPOLOGY: linear</p> <p>MOLECULE TYPE: protein</p> <p>US-08-589-756-3</p>			
QY	169	LKSIINAPRGKNDGCGMYISNIDTGTDRHKAMRIDDDAKASMRKKEDLKGTDKNY----	245
Db	111	YKTLDEKKGK--GAGTVVAVIDAGFDKKNHEAMRLTDKAKARYQ-SKEDLEKAKKEHGITT	166

```

OY      246  ---WMSDKIPAHFNYNGKITEVEYDDGRDYFDPDHNGHILAGINTEODIKNFNCID 302
Db      167  YGEWVNDVAVYIYHDSKDKTAIVDQ-----EHTHVSIGSLGN-APSETKPEPYLE 216
OY      303  GIAPNAQJFFSYKM-----YSDAGSGFAGDETFMHAIEDSIKHNVDVSVSGFTGTGLVG 357
Db      217  GAMPBAQQLLMKRELIVNGIADARNYA-----QAIRDAINIGAKVIMNSFGNALAYAN 270
OY      358  --EKYQAIRALKRKAGITPMVATGNVATISASSSSMDLVANNHLKMTDPGNGVTRTAHEDA 415
Db      271  LPDEFKKAFEDYAKSGVSIYTSAGNDSFEGCKTRLPLA--DH---PDGVGVTPPAAAST 325
OY      416  IAVA--SAKNQTVEDDKVNIGESFEKYRNIGAFPKSKITTEDEGTAKPSK-LKEVYIGK 472
Db      326  LTVASYSPDKQULETVYRKKTADQOK-----EMPLSTNR---FEPNKAYDVAAYNR 374
OY      473  GODODLIGLGRKIAVMDRITYTKDLNAFKKAMDKGARAIWVNTVYNYND----- 525
Db      375  GTKEDDF--KQVYKGIATLER--GDIDFKDKIAKAKAGAVGLI-----YDNODKGFIEL 427
OY      526  -NMWTLPLPMGVEADGTSQVFSISGDDGVKIKMNNINPDKTEYVRNNKEDPKKLEQY 584
Db      428  PNVQDMPAA-----FISKQGLLKD--NP----- 450
OY      585  PIDMESFNKNPNVGEKEIDFKFAPDIDKELYKDIIYFAPS-----TSMGPRIDL 637
Db      451  -----QXTITFNATPK-----VLPTASGTRKLSRFSMWGLTADGN 484
OY      638  LKPDVSAPEGNIKISTLANYNCK 659
Db      485  IKPDIAAPGQDILS--SVANNK 504

RESULT      5
US-09-206-800-3
; Sequence 3, Application US/09206800
; Patent No. 6270775
; GENERAL INFORMATION:
;   APPLICANT: Cleary, P. P.
;   TITLE OF INVENTION: Streptococcal C5a peptidase vaccine
;   FILE REFERENCE: 600.349US2
;   CURRENT APPLICATION NUMBER: US/09/206, 800
;   CURRENT FILING DATE: 1998-12-07
;   EARLIER APPLICATION NUMBER: US 08/589, 756
;   EARLIER FILING DATE: 1996-01-22
;   NUMBER OF SEQ ID NOS: 11
;   SOFTWARE: FastSeq for Windows Version 3.0
;   SEQ ID NO 3
;   LENGTH: 1150
;   TYPE: PRF
;   ORGANISM: Streptococcus agalactiae
US-09-206-800-3

Query Match      8.6%; Score 296.5; DB 4; Length 1150;
Best Local Similarity 24.3%; Pred. No. 9,4e-12;
Matches 122; Conservative 78; Mismatches 163; Indels 139; Gaps 24.

OY      189  LKSINAPFGKNEFDGRGWIYISNIDGTDIRKHKAMRIDDDAKASMRKKEDLKGTDKNY--- 245
Db      111  VTLTQEKAG--GAGTVAIVADAGFDKNHEMARLTDKTKARYQ--SKEDLEKAKKEHGIT 166
OY      246  ---WISDIPAHFNYNGKITVEYDDGRDYFDPDHNGHILAGINTEODIKNFNCID 302
Db      167  YGEWVNDVAVYIYHDSKDKTAIVDQ-----EHTHVSIGSLGN-APSETKPEPYLE 216
OY      303  GIAPNAQJFFSYKM-----YSDAGSGFAGDETFMHAIEDSIKHNVDVSVSGFTGTGLVG 357
Db      217  GAMPBAQQLLMKRELIVNGIADARNYA-----QAIRDAINIGAKVIMNSFGNALAYAN 270
OY      358  --EKYQAIRALKRKAGITPMVATGNVATISASSSSMDLVANNHLKMTDPGNGVTRTAHEDA 415
Db      271  LPDEFKKAFEDYAKSGVSIYTSAGNDSFEGCKTRLPLA--DH---PDGVGVTPPAAAST 325
OY      416  IAVA--SAKNQTVEDDKVNIGESFEKYRNIGAFPKSKITTEDEGTAKPSK-LKEVYIGK 472
Db      326  LTVASYSPDKQULETVYRKKTADQOK-----EMPLSTNR---FEPNKAYDVAAYNR 374
OY      473  GODODLIGLGRKIAVMDRITYTKDLNAFKKAMDKGARAIWVNTVYNYND----- 525
Db      375  GTKEDDF--KQVYKGIATLER--GDIDFKDKIAKAKAGAVGLI-----YDNODKGFIEL 427
OY      526  -NMWTLPLPMGVEADGTSQVFSISGDDGVKIKMNNINPDKTEYVRNNKEDPKKLEQY 584
Db      428  PNVQDMPAA-----FISKQGLLKD--NP----- 450
OY      585  PIDMESFNKNPNVGEKEIDFKFAPDIDKELYKDIIYFAPS-----TSMGPRIDL 637
Db      451  -----QXTITFNATPK-----VLPTASGTRKLSRFSMWGLTADGN 484
OY      638  LKPDVSAPEGNIKISTLANYNCK 659
Db      485  IKPDIAAPGQDILS--SVANNK 504

358  --EKYQAIRALKRKAGITPMVATGNVATISASSSSMDLVANNHLKMTDPGNGVTRTAHEDA 415

```

Db 271 LPDETKAFDYAKSGVSVTSAGNDSSFGKTRPLA--DH---PDYGVGTGTPAAADST 325
QY 416 IAVA--SAKNOTVERFDKYNIGSESPKRYNIGAFDPKSKITTNEDGTAKPSK-LKFVYIK 472
Db 326 LTVASYSFDKQLETETVRKVTADQDQK-----EMPVLSTNR---FEPNKAIDYAVANR 374
QY 473 GQDDDLGLDRLGKIAVMDRIYTKDLKNAFKKAMDGRARAIVVNTVNYNRD----- 525
Db 375 GKEDDF-KDYKGIKIALIER-GDIDFKDKIAKAKAGAVGLI-----YDNDKGFIEL 427
QY 526 -NMTELPAMGYEADGEGTSQVFSISGDGCVKLMNMINPDKTEVRKNNKEDFKLEQY 584
Db 428 PNVQDMPAA-----FISRKDGLLKD--NP----- 450
QY 585 PIDMESFNKNPNVGEDEKIDFKFAPDDDKELKEDIIVPAGS-----TSGPRIDL 637
Db 451 -----QKTIFFNATPK-----VLPTASGTKLSRFSMGLTADGN 484
QY 638 LKPDVSAPGKNIKSTPLNYNGK 659
b 485 IKPDIAFGDILS--SVANNK 504

RESULT 6
US-08-589-756-2
; Sequence 2, Application US/08589756
; Patent No. 5846547
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: STREPTOCOCCAL C5a PEPTIDASE VACCINE
; NUMBER OF SEQUENCES: 5
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,756
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1167 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-589-756-2

Query Match 8.5%; Score 293; DB 2; Length 1167;
Best Local Similarity 23.0%; Pred. No. 1.6e-11;
Matches 139; Conservative 93; Mismatches 207; Indels 166; Gaps 29;

QY 97 NPKKEDKVVYIAEFKDKSEGEKAI-----KELSLKNTKVLTYDRIFNGSAIETT 147
Db 24 NAQSDIKANTYTE--DTPVTEQAVETPOPTAVSEEVSSKETKTPOTPD-----DAEETI 76
QY 148 PDLNDKIKQIEGISSVERAKQVOPMNNHARKEIGVEAID--YLSINAPFGKNFGDGRM 205
Db 77 ADDANDLAPQAPAKTAD-----TPATSKATIRDLNDPSQVKTLOEKAGK---GAGT 124
QY 206 VISINDTGTDRHKARIDDDAKASMRKKEDLKGTDKNY-----WLSDKIPHAFNYYN 259
Db 125 VVAVIDAGFDKNEHAWRLDTKTARYQ--SKEDLEKAKKEHGIYGEVWVNDKVAAYHYDYSK 183
QY 260 GCKITVEKYDDGGRDYPDHGHIAGILAGNTEODIKNFNGIDGIAFNAQIFSYKM---- 315
Db 184 DKGTAVDQ-----EHGTHVSGILSGN-APSETEKPEYRLGAMPEAQDLLMRVEIYN 233
QY 316 -YSDAGSGFAGDETFMHAIEDSIKHNVDVSVSSGFTGTGLVG--EKYQAIIRALKRAGI 372
Db 234 GLADYARNYA-----QAIRDAVNLGAKVYINNSFGNALALAYANLDETKKADYAKSKCV 287
QY 373 PMVYATGNVATSSASSSMDLVANNHLKMTDTGNVTRTAHEDAIAVA--SAKNOTVERFDK 430

Db 288 SIYTSAGNDSSFGKTRPLA--DH---PDYGVGTGTPAAADSTLTVASYSFDKQLETETAM 342
QY 431 VNIGSESPKRYNIGAFDPKSKITTNEDGTAKPSK-LKFVYIGGQDDDLGLDRLGKIAV 489
Db 343 VKTDDQDQK-----EMPVLSTNR---FEPNKAIDYAVANRKKEDDF-KDYKGIKIAL 390
QY 490 MDRITYTKDLKNAFKKAMDGRARAIVVNTVNYNRD-----NMTELPAMGYEADGEGT 541
Db 391 IER-GDIDFKDKYANNAKAGAVGLI-----YDNDKGFIELPNVQMPAA----- 436
QY 542 KSGVFSISGDGCVKLMNMINPDKTEVRKNNKEDFKLEQYYPIDMESFNKNPNVGE 601
Db 437 -----FISRKDGLLKD--NP-----Q 451
QY 602 KETDFKAPDDDKELKEDIIVPAGS-----TSGPRIDLKPDVSAPGKNIKSTLN 654
Db 452 KTIFFNATPK-----VLPTASGTKLSRFSMGLTADGNIKPDIAFGDILS--S 499
QY 655 VINGK 659
Db 500 VANNK 504

RESULT 7
US-09-206-800-2
; Sequence 2, Application US/09206800
; Patent No. 6270775
; GENERAL INFORMATION:
; APPLICANT: Cleary, P. P.
; TITLE OF INVENTION: Streptococcal C5a peptidase vaccine
; FILE REFERENCE: 600.349US2
; CURRENT APPLICATION NUMBER: US/09/206,800
; EARLIER APPLICATION NUMBER: US 08/589,756
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes, strain 12
; US-09-206-800-2

Query Match 8.5%; Score 293; DB 4; Length 1167;
Best Local Similarity 23.0%; Pred. No. 1.6e-11;
Matches 139; Conservative 93; Mismatches 207; Indels 166; Gaps 29;

QY 97 NPKKEDKVVYIAEFKDKSEGEKAI-----KELSLKNTKVLTYDRIFNGSAIETT 147
Db 24 NAQSDIKANTYTE--DTPVTEQAVETPOPTAVSEEVSSKETKTPOTPD-----DAEETI 76
QY 148 PDLNDKIKQIEGISSVERAKQVOPMNNHARKEIGVEAID--YLSINAPFGKNFGDGRM 205
Db 77 ADDANDLAPQAPAKTAD-----TPATSKATIRDLNDPSQVKTLOEKAGK---GAGT 124
QY 206 VISINDTGTDRHKARIDDDAKASMRKKEDLKGTDKNY-----WLSDKIPHAFNYYN 259
Db 125 VVAVIDAGFDKNEHAWRLDTKTARYQ--SKEDLEKAKKEHGIYGEVWVNDKVAAYHYDYSK 183
QY 260 GCKITVEKYDDGGRDYPDHGHIAGILAGNTEODIKNFNGIDGIAFNAQIFSYKM---- 315
Db 184 DKGTAVDQ-----EHGTHVSGILSGN-APSETEKPEYRLGAMPEAQDLLMRVEIYN 233
QY 316 -YSDAGSGFAGDETFMHAIEDSIKHNVDVSVSSGFTGTGLVG--EKYQAIIRALKRAGI 372
Db 234 GLADYARNYA-----QAIRDAVNLGAKVYINNSFGNALALAYANLDETKKAFDYAKSKGV 287
QY 373 PMVYATGNVATSSASSSMDLVANNHLKMTDTGNVTRTAHEDAIAVA--SAKNOTVERFDK 430
Db 288 SIYTSAGNDSSFGKTRPLA--DH---PDYGVGTGTPAAADSTLTVASYSFDKQLETETAM 342


```

QY 431 VNIIGSEFKRNIGAFEDSKITTTNEDGTAKPSK-LKFVYIGGODDILGLDRKIAV 489
DB 343 VKTDDODK-----EMVPLSTNR---FEPNKAIDYAYANRGMKEDD-KDYKRIAL 390
QY 490 MDRIYRKDLKNAFKKAMDGAIRAIWNTVNYNRD-----NMTELPAWGEADEGT 541
DB 391 IER-GDIDFKDKAANKKACAVGLI-----YDNDKGFIELPNDQMPAA----- 436
QY 542 KSGVFSISGDDGYKLVNMINPDKTEVKRNKEDFKDLQYYPIDMESFNKPNVGE 601
DB 437 -----FISRKDGLLKD--NP-----Q 451
QY 602 KEIDFAPPTDKELYKEDIIVPAGS-----TSMGPRIDLLKPDVSAFGKNISTLN 654
DB 452 KITTFNATPK-----VLPTASGTRLSRFSWGLTADGNIKPIDIAAPGODILS--S 499
QY 655 VINGK 659
DB 500 VANNK 504

RESULT 8
US-09-206-800-8
; Sequence 8, Application US/09206800
; Patent No. 6270775
; GENERAL INFORMATION:
; APPLICANT: Cleary, P. P.
; TITLE OF INVENTION: Streptococcal C5a peptidase vaccine
; FILE REFERENCE: 600.349052
; CURRENT APPLICATION NUMBER: US/09/206,800
; EARLIER FILING DATE: 1998-12-07
; EARLIER APPLICATION NUMBER: US 08/589,756
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 8
; LENGTH: 383
; TYPE: PRF
; ORGANISM: Streptococcus agalactiae
US-09-206-800-8

Query Match 7.9%; Score 273; DB 4; Length 383;
Best Local Similarity 24.0%; Pred. No. 7.9e-11;
Matches 115; Conservative 73; Mismatches 156; Indels 136; Gaps 23;

QY 211 DTGTDRHKAMRIDDAKASMRKKEDLKTGDKNY-----WLSDKIPAHFNYNGKIT 264
DB 1 DAGFDKNHEAMRLTDKTKARYQ--SKEDLEKAKKEHGITYGEWVNDKVAAYHYHDSKDKTA 59
QY 265 VEKYDDGRDYFDPHGMHIAIAGNTEODIKNFNGIDGIAPNNAQIFSYKM-----YSDA 319
DB 60 VDO-----EHGTVSGILSGN--APSETKEPRLEGAMPQQLMLMVEIYNGLADY 109
QY 320 GSGFAGDETFMHAIEDSIKHNVDVSVSGFTGTGLVG--EKYQOALRALRKAGIPVVA 377
DB 110 ARNYA-----QAIRDAINIGAKVYINSGFNAALAYANLPDETKKADYAKSGVSLVTS 163
QY 378 TGNVATSSASSMDLVANNLKMKTGTGNVTRTAHEDAIAVA--SAKNQVVEFDKVNIGG 435
DB 164 AGNDSSRGKTRPLA--DH---PDYGVGTTPAAADSTLVASSPKQLTEYRVRTAD 218
QY 436 ESFRYRNIGAFPDKSKITTTNEDGTAKPSK-LKFVYIGGODDILGLDRKIAVMDRIY 494
DB 219 QODK-----EMVPLSTNR---FEPNKAIDYAYANRGMKEDD-KDYKRIALIER-G 265
QY 495 TKDLKNAFKKAMDGAIRAIWNTVNYNRD-----NMTELPAWGEADEGTQOVF 546
DB 266 DIFDKDIAAKKAGAVGLI-----YDNDKGFIELPNDQMPAA----- 307
QY 547 SISGDDGVKLMNINPDKTEVKRNKEDFKDLQYYPIDMESFNKPNVGEKEIDF 606
DB 308 FISRKDGLLKD--NP-----QKITTF 327
```

```

QY 607 KFAPDVDELYKEDIIVPAGS-----TSMGPRIDLLKPDVSAFGKNISTLVINGK 659
DB 328 NATPK-----VLPTASGTRLSRFSWGLTADGNIKPIDIAAPGODILS--SVANNK 375

RESULT 9
US-09-206-800-11
; Sequence 11, Application US/09206800
; Patent No. 6270775
; GENERAL INFORMATION:
; APPLICANT: Cleary, P. P.
; TITLE OF INVENTION: Streptococcal C5a peptidase vaccine
; FILE REFERENCE: 600.349052
; CURRENT APPLICATION NUMBER: US/09/206,800
; EARLIER FILING DATE: 1998-12-07
; EARLIER APPLICATION NUMBER: US 08/589,756
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 11
; LENGTH: 921
; TYPE: PRF
; ORGANISM: Streptococcus agalactiae
US-09-206-800-11

Query Match 7.9%; Score 273; DB 4; Length 921;
Best Local Similarity 24.0%; Pred. No. 2.5e-10;
Matches 115; Conservative 73; Mismatches 156; Indels 136; Gaps 23;

QY 211 DTGTDRHKAMRIDDAKASMRKKEDLKTGDKNY-----WLSDKIPAHFNYNGKIT 264
DB 1 DAGFDKNHEAMRLTDKTKARYQ--SKEDLEKAKKEHGITYGEWVNDKVAAYHYHDSKDKTA 59
QY 265 VEKYDDGRDYFDPHGMHIAIAGNTEODIKNFNGIDGIAPNNAQIFSYKM-----YSDA 319
DB 60 VDO-----EHGTVSGILSGN--APSETKEPRLEGAMPQQLMLMVEIYNGLADY 109
QY 320 GSGFAGDETFMHAIEDSIKHNVDVSVSGFTGTGLVG--EKYQOALRALRKAGIPVVA 377
DB 110 ARNYA-----QAIRDAINIGAKVYINSGFNAALAYANLPDETKKADYAKSGVSLVTS 163
QY 378 TGNVATSSASSMDLVANNLKMKTGTGNVTRTAHEDAIAVA--SAKNQVVEFDKVNIGG 435
DB 164 AGNDSSRGKTRPLA--DH---PDYGVGTTPAAADSTLVASSPKQLTEYRVRTAD 218
QY 436 ESFRYRNIGAFPDKSKITTTNEDGTAKPSK-LKFVYIGGODDILGLDRKIAVMDRIY 494
DB 219 QODK-----EMVPLSTNR---FEPNKAIDYAYANRGMKEDD-KDYKRIALIER-G 265
QY 495 TKDLKNAFKKAMDGAIRAIWNTVNYNRD-----NMTELPAWGEADEGTQOVF 546
DB 266 DIFDKDIAAKKAGAVGLI-----YDNDKGFIELPNDQMPAA----- 307
QY 547 SISGDDGVKLMNINPDKTEVKRNKEDFKDLQYYPIDMESFNKPNVGEKEIDF 606
DB 308 FISRKDGLLKD--NP-----QKITTF 327
QY 607 KFAPDVDELYKEDIIVPAGS-----TSMGPRIDLLKPDVSAFGKNISTLVINGK 659
DB 328 NATPK-----VLPTASGTRLSRFSWGLTADGNIKPIDIAAPGODILS--SVANNK 375

RESULT 10
US-09-206-800-6
; Sequence 6, Application US/09206800
; Patent No. 6270775
; GENERAL INFORMATION:
; APPLICANT: Cleary, P. P.
; TITLE OF INVENTION: Streptococcal C5a peptidase vaccine
; FILE REFERENCE: 600.349052
; CURRENT APPLICATION NUMBER: US/09/206,800
```

```

: CURRENT FILING DATE: 1998-12-07
: EARLIER APPLICATION NUMBER: US 08/589,756
: EARLIER FILING DATE: 1996-01-22
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO: 6
: LENGTH: 383
: TYPE: PRT
: ORGANISM: Streptococcus pyogenes, strain 49
US-09-206-800-6

```

```

Query Match      7.8%; Score 268; DB 4; Length 383;
Best Local Similarity 24.7%; Pred. No. 1,76-10;
Matches 118; Conservative 69; Mismatches 148; Indels 142; Gaps 25;

```

```

QY 211 DTGTDYRHKAMRIDDDAKASRFKKEDLKGTDKNY-----WLSKIPHAENYNGKAIT 264
DB 1 DAGFDKNHEAMRLTDKAKARYO-SKEDLEKAKKEGITYGEMVNDKAVYHYDYSKDGKTA 59
QY 265 VEKYDDGDRYDPHGMHAGILAGNDTEQDIKNFNGIDGAPNMOIFSYK-----YSDA 319
DB 60 VDO-----EHTHVSGLISGN-APSETKEPYRLEGAMPEAOQLLMRVEIYNGGLADY 109
QY 320 GSGFAGDETMFHAIEDSIKHNVDVVSSG---FTGTGLVEKYQAIIRALKRAGIPMYV 376
DB 110 ARNYA-----QAIRDAVNLGAKYINMSFGNALAYANLPDETKKPFYAKSK-GVRIYV 162
QY 377 ATGNVATSSSSMDLVANNHLKMTDGNVTRTAHEDAIAVA--SAKQTVPEFDKVNIG 434
DB 163 TAGNDSSFGKTRPLA--DH---PDYGVGTAPAAADSTLVVASYSPNOLETAMVATD 217
QY 435 GESFERNIGAFEDSKITTTNEDGKAPSK-LKFVYICKGDDQDILGLDGLKGTIAVMRI 493
DB 218 DQDK-----EMPVLTNR---FEPNKAVDYANRGMKEDDF-KDVKGKIALIERS 265
QY 494 ---YTKDLKNAFKRAMDGARAIMVNVNYYNRD-----NMTLPAWGYEADDECT 542
DB 266 DIDFDKIANA-KKA---GAVGVLT---YDNQDKGFIELPNVDQMPAA----- 307
QY 543 SQVESISGDDGVKLMNMINPDKTEVRKNNKEDFKDLQYYPIDMESFNSKNKPVGDEK 602
DB 308 ---FISRKDGLL-----KD-----NSQK 323
QY 603 EIDKFAPDTEKELYKEDIIVPAGS-----TSWGPRIIDLLKPDVAPGKNIKST 652
DB 324 TITFNATPK-----VLPASGTRLSRFSWGTLADGNIKPDIAAPGODILSS 370

```

```

RESULT 11
US-09-206-800-9
: Sequence 9, Application US/09206800
: Patent No. 6270775
: GENERAL INFORMATION:
: APPLICANT: Cleary, P. P.
: TITLE OF INVENTION: Streptococcal C5a peptidase vaccine
: FILE REFERENCE: 600.349052
: CURRENT APPLICATION NUMBER: US/09/206,800
: CURRENT FILING DATE: 1998-12-07
: EARLIER APPLICATION NUMBER: US 08/589,756
: EARLIER FILING DATE: 1996-01-22
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 9
: LENGTH: 921
: TYPE: PRT
: ORGANISM: Streptococcus pyogenes, strain 49
US-09-206-800-9

```

```

Query Match      7.8%; Score 268; DB 4; Length 921;
Best Local Similarity 24.7%; Pred. No. 5,56-10;
Matches 118; Conservative 69; Mismatches 148; Indels 142; Gaps 25;

```

```

QY 211 DTGTDYRHKAMRIDDDAKASRFKKEDLKGTDKNY-----WLSKIPHAENYNGKAIT 264
DB 1 DAGFDKNHEAMRLTDKAKARYO-SKEDLEKAKKEGITYGEMVNDKAVYHYDYSKDGKTA 59
QY 265 VEKYDDGDRYDPHGMHAGILAGNDTEQDIKNFNGIDGAPNMOIFSYK-----YSDA 319
DB 60 VDO-----EHTHVSGLISGN-APSETKEPYRLEGAMPEAOQLLMRVEIYNGGLADY 109
QY 320 GSGFAGDETMFHAIEDSIKHNVDVVSSG---FTGTGLVEKYQAIIRALKRAGIPMYV 376
DB 110 ARNYA-----QAIRDAVNLGAKYINMSFGNALAYANLPDETKKPFYAKSK-GVRIYV 162
QY 377 ATGNVATSSSSMDLVANNHLKMTDGNVTRTAHEDAIAVA--SAKQTVPEFDKVNIG 434
DB 163 TAGNDSSFGKTRPLA--DH---PDYGVGTAPAAADSTLVVASYSPNOLETAMVATD 217
QY 435 GESFERNIGAFEDSKITTTNEDGKAPSK-LKFVYICKGDDQDILGLDGLKGTIAVMRI 493
DB 218 DQDK-----EMPVLTNR---FEPNKAVDYANRGMKEDDF-KDVKGKIALIERS 265
QY 494 ---YTKDLKNAFKRAMDGARAIMVNVNYYNRD-----NMTLPAWGYEADDECT 542
DB 266 DIDFDKIANA-KKA---GAVGVLT---YDNQDKGFIELPNVDQMPAA----- 307
QY 543 SQVESISGDDGVKLMNMINPDKTEVRKNNKEDFKDLQYYPIDMESFNSKNKPVGDEK 602
DB 308 ---FISRKDGLL-----KD-----NSQK 323
QY 603 EIDKFAPDTEKELYKEDIIVPAGS-----TSWGPRIIDLLKPDVAPGKNIKST 652
DB 324 TITFNATPK-----VLPASGTRLSRFSWGTLADGNIKPDIAAPGODILSS 370

```

```

RESULT 12
US-09-206-800-7
: Sequence 7, Application US/09206800
: Patent No. 6270775
: GENERAL INFORMATION:
: APPLICANT: Cleary, P. P.
: TITLE OF INVENTION: Streptococcal C5a peptidase vaccine
: FILE REFERENCE: 600.349052
: CURRENT APPLICATION NUMBER: US/09/206,800
: CURRENT FILING DATE: 1998-12-07
: EARLIER APPLICATION NUMBER: US 08/589,756
: EARLIER FILING DATE: 1996-01-22
: NUMBER OF SEQ ID NOS: 11
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO: 7
: LENGTH: 383
: TYPE: PRT
: ORGANISM: Streptococcus pyogenes, strain 12
US-09-206-800-7

```

```

Query Match      7.6%; Score 264; DB 4; Length 383;
Best Local Similarity 23.5%; Pred. No. 3,16-10;
Matches 113; Conservative 73; Mismatches 158; Indels 136; Gaps 23;

```

```

QY 211 DTGTDYRHKAMRIDDDAKASRFKKEDLKGTDKNY-----WLSKIPHAENYNGKAIT 264
DB 1 DAGFDKNHEAMRLTDKAKARYO-SKEDLEKAKKEGITYGEMVNDKAVYHYDYSKDGKTA 59
QY 265 VEKYDDGDRYDPHGMHAGILAGNDTEQDIKNFNGIDGAPNMOIFSYK-----YSDA 319
DB 60 VDO-----EHTHVSGLISGN-APSETKEPYRLEGAMPEAOQLLMRVEIYNGGLADY 109
QY 320 GSGFAGDETMFHAIEDSIKHNVDVVSSGFTGLVNG--EKYQAIIRALKRAGIPMYV 377
DB 110 ARNYA-----QAIRDAVNLGAKYINMSFGNALAYANLPDETKKPFYAKSKGSIYTS 163
QY 378 TGNVATSSSSMDLVANNHLKMTDGNVTRTAHEDAIAVA--SAKQTVPEFDKVNIG 435

```

Db 164 AGDSSFGKTRPLA--DH---PDYGVGTAAADSTLTVAISPDKOLTEETAMKTDD 218
 QY 436 ESEKYRNIGAFEDSKITTNEDGTAPSK-LKEVYIGKGODDLGLDGRKIAVMDRIY 494
 Db 219 QQDK-----EMPVLSNR---FEPNKAIDYAVANRGKEDDF-KDVKGKIALIER-G 265
 QY 495 TYDLKAFFKAMDKARAIWNTVTVYNRD-----NMTLEPAMGYEADGTSQYV 546
 Db 266 DIDEKDVANAKAGAVGLI-----YDNODKGFPIELPVDQMPAA----- 307
 QY 547 STSGDGVLMNINPDKTEVRNKKDEKDLQYPIIDMSFNSNKNPVGDEKEIDF 606
 Db 308 FLSKDGLLKD--NP-----QKTTTF 327
 QY 607 KFAPDDEKELYEDIVPAGS-----TSWGPRIIDLKPDVAPGKNIKSTLYNKG 659
 Db 328 NATPK-----VLPRASGTRKLSRFSSWGLTADGNIKPDIAAGQDILS--SVANNK 375

SULT 13
 -09-206-800-10
 Sequence 10, Application US/09206800
 Patent No. 6270775
 GENERAL INFORMATION:
 APPLICANT: Cleary, P. P.
 TITLE OF INVENTION: Streptococcal C5a peptidase vaccine
 FILE REFERENCE: 600.349US2
 CURRENT APPLICATION NUMBER: US/09/206,800
 EARLIER FILING DATE: 1998-12-07
 CURRENT APPLICATION NUMBER: US 08/589,756
 EARLIER FILING DATE: 1996-01-22
 NUMBER OF SEQ ID NOS: 11
 SOFTWARE: FastSeq for Windows Version 3.0
 SEQ ID NO 10
 LENGTH: 921
 TYPE: PRF
 ORGANISM: Streptococcus pyogenes, strain 12
 US-09-206-800-10

Query Match 7.6%; Score 264; DB 4; Length 921;
 Best Local Similarity 23.5%; Pred. No. 1e-09;
 Matches 113; Conservative 73; Mismatches 158; Indels 136; Gaps 23;
 QY 211 DTGTDVRHKAMRIDDAKASMRKEDLKTGDKNY-----WLSDKRPHAFNYNNGKIT 264
 Db 1 DAQFDKNHEWRLTDKTKARYQ-SKEDLEKAKKEHGTTEWYNDKVAIYHDSKDEKTA 59
 QY 265 VEKYDGRDYFDPGHMIIAGLNDTEODIKNFNGIDGIAFPAQJFSYKM-----YSDA 319
 Db 60 VDO-----EHGTHVSGILSGN-APSETKEPYRLEGAMPEAQLLMVEIYNGLDY 109
 QY 320 GSGFADDETFMFAIDSIKINVDYVSSSGFTGTGLVG--EKYQAIKALRKAGIPVYA 377
 Db 110 ARNYA-----QAIRAVNLGAVYINNSFGNAALAYANLPDETKKADYAKSGVSVTS 163
 QY 378 TGNVATSSSSMDLYANHLKMTDGNVTRTAHEDAIVA--SAKNQTFEFDKYNIGG 435
 Db 164 AGNDSSTSGCTRIPLA--DH---PDYGVGTAAADSTLTVAISPDKOLTEETAMKTDD 218
 QY 436 ESEKYRNIGAFEDSKITTNEDGTAPSK-LKEVYIGKGODDLGLDGRKIAVMDRIY 494
 Db 219 QQDK-----EMPVLSNR---FEPNKAIDYAVANRGKEDDF-KDVKGKIALIER-G 265
 QY 495 TYDLKAFFKAMDKARAIWNTVTVYNRD-----NMTLEPAMGYEADGTSQYV 546
 Db 266 DIDEKDVANAKAGAVGLI-----YDNODKGFPIELPVDQMPAA----- 307
 QY 547 STSGDGVLMNINPDKTEVRNKKDEKDLQYPIIDMSFNSNKNPVGDEKEIDF 606
 Db 308 FLSKDGLLKD--NP-----QKTTTF 327
 QY 607 KFAPDDEKELYEDIVPAGS-----TSWGPRIIDLKPDVAPGKNIKSTLYNKG 659

Db 328 NATPK-----VLPRASGTRKLSRFSSWGLTADGNIKPDIAAGQDILS--SVANNK 375

RESULT 14
 US-08-894-818B-35
 Sequence 35, Application US/08894818B
 Patent No. 6261822
 GENERAL INFORMATION:
 APPLICANT: TAKAKURA, Hikaru
 APPLICANT: MORISHITA, Mio
 APPLICANT: YAMAMOTO, Katsuhiko
 APPLICANT: MITTA, Masanori
 APPLICANT: ASADA, Kiyozo
 APPLICANT: TSUNASAWA, Susumu
 APPLICANT: KATO, Ikunoshin
 TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Browdy and Neimark
 STREET: 419 Seventh Street N.W., Ste. 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: United States of America
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/894,818B
 FILING DATE: 20-MAY-1998
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP96/03253
 FILING DATE: 07-NOV-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 33285/1995
 FILING DATE: 12-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Browdy, Roger L.
 REGISTRATION NUMBER: 25,618
 REFERENCE/DOCKET NUMBER: TAKAKURA-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-5197
 TELEFAX: (202) 737-3528
 INFORMATION FOR SEQ ID NO: 35:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 654 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-894-818B-35

Query Match 6.8%; Score 236; DB 4; Length 654;
 Best Local Similarity 22.3%; Pred. No. 4.6e-08;
 Matches 144; Conservative 91; Mismatches 252; Indels 158; Gaps 29;
 QY 62 VIDNNTSNEAKIKRE-ENSKRSQGDYTDSPFNKNTN-NPKKEDKVVYIAEFKKEGSEKA 119
 Db 17 VVGSAVAAPKEKVEQYRNVKNGYGLTLPGLFRKIQKLNPREISTVYV--FENHREKEIA 74
 QY 120 IKELSSIKNTKYLTYDRINGSAIETTPNDLKIKQIEG-----ISSVEAOKVQPPAMNH 175
 Db 75 VRVL-ELMGAKVRXY-HIIPAIADLKVDLLVIGLGGKAKLSGVRFIOEDYKTVS 132
 QY 176 ARKEIGVEAIDYKLSINAPFGKNFDGRGAVISNIDPTGTDYRHKAMRIDDAKASMRFK 235
 Db 133 AELLE-GIDESAAQVMATYV-WNLGYDSGITIGIIDTGIDASH----- 173

QY 236 EDLKTGDKNWYLSKIPHAFFNYNGKITVEKYYDDGRDPFPHGMHLAGNDTEODI 295
 Db 174 PDLOG-----KVGWDFVNGRSY---PYDD-----HGCHTHASIAAGGASNG 216
 QY 296 KFNFGIDGAPNAOIFSYKMSDAGSG-----FAGDETFHAIEDSKHNVDVSVSSGF 350
 Db 217 K-----YKMGAPGAKLACIKVLAGDAGSGSISITIKGVE---WAVDNKDKYGIKVINLSIGS 269
 QY 351 TGTGLVGEKYMQAIRALKKAGIPMVVATGN-----YANSA 385
 Db 270 SOSSDGTDALSOAVNAADAGLVVVAAGNSGPNKYYTIGSPAASKVITGAVDKYDVT 329
 QY 386 SSSSMDLVANNHLK--MTDGNVTRTAHEDALIVASAKNOTVEDEKYNIGSESEFYKNI 443
 Db 330 SFSRSGPTADRLKPEYVAPGNWIIAAR-----ASGTSMGQPINDYTAAGTSMATPHV 384
 QY 444 GAFEDKSTTTNEDGTAPSKLFYVYIGKODDOLIGDLGKTAIVMDRITTKDKNAFK 503
 Db 385 AGI---MALLQAHPSWTPDKVKTALI---ETADIVKPD-----EIDAIYAGAGRVNAFK 433
 QY 504 KAMDGARAIVVNTVNYNRDNTLPAWGEDEGTSQVFSISGDDGVK---LWMMI 560
 Db 434 -----AINT---DNIAKLVTGTIVANKSGQTHQFVISCASEVTATLYMDNA 476
 QY 561 NPKRTEV--KRNKEDPK-----DKLEQY-----YPIDMESFN 592
 Db 477 NSDDLXYLDXNGNOVDYSYRAYYGFEKVGYNPTDGTWIKIVVSYSGSANVQYDVVSDG 536
 QY 593 S-----KRPVNGDEKEIDFKFAPDTEKELYED-----IIVAGST 628
 Db 537 SLSPGSSPSQPEPTVDAKTEGSDHYHYDRSDPTFTMTVNSGAT 591

RESULT 15
 US-08-894-818B-1
 Sequence 1, Application US/08894818B
 Patent No. 6261822
 GENERAL INFORMATION:
 APPLICANT: TAKAKURA, Hikaru
 APPLICANT: MORISHITA, Mio
 APPLICANT: YAMAMOTO, Katsuhiko
 APPLICANT: MITTA, Masanori
 APPLICANT: ASADA, Kiyozo
 APPLICANT: TSUNASAWA, Susumu
 APPLICANT: KATO, Ikunoshi
 TITLE OF INVENTION: HYPERTHERMOSTABLE PROTEASE GENES
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Browdy and Neimark
 STREET: 419 Seventh Street N.W., Ste. 300
 CITY: Washington
 STATE: D.C.
 COUNTRY: United States of America
 ZIP: 20004
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/894,818B
 FILING DATE: 20-MAY-1998
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: PCT/JP96/03253
 FILING DATE: 07-NOV-1996
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 323285/1995
 FILING DATE: 12-DEC-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Browdy, Roger L.
 REGISTRATION NUMBER: 25, 618
 REFERENCE/DOCKET NUMBER: TAKAKURA-1

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 628-5197
 TELEFAX: (202) 737-3528
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 659 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-894-818B-1

Query Match 6.6%; Score 229.5; DB 4; Length 659;
 Best Local Similarity 21.8%; Pred. No. 1,36-07;
 Matches 135; Conservative 80; Mismatches 239; Indels 157; Gaps 27;

QY 61 PYIDNNTSNEEAKIKENSNGSGDYDTSFVNK--NTENPKKE-DKVYIAEFKKESEGEK 118
 Db 28 PVRNNNAVOQ-----KNYGLTFGLFKKYQRMNMNOEVDVIMFGSGYGRDRAVK 77
 QY 119 AIKELSLKNTKVLTYDRIENGSAIETTPDNLKI-----KQIEGISSVERAQ 167
 Db 78 VLR-----LMGAOVKYSY-KIIPAVAVKIKARDLLIAGMIDTGYFGNTRVSGIFIOEDY 132
 QY 168 KVOPEMMNHARKIEVEEAIDYLSKINAFKNGDGRGVNISIDTGYRHRKAMRIDDA 227
 Db 133 KYQYDDATSVSQIGADYWNLS-----GYDGSQVVAIVDTGIDANH----- 174
 QY 228 KASRFEKKEDIKGTDKNWLSDKIPHAFFNYNGKITVEKYDDGRDPFPHGMHLAGLA 287
 Db 175 -----PDLKG-----KVIGWYDAVNGRS---TPYDD-----QCHGTHVAGIVA 209
 QY 288 GNDT--EODIKFNFGIOGSIAPNAOIFSYKMSDAGSG-----FAGDETFHAIEDSKHN 340
 Db 210 GTGSVNSQYI-----GVAFGAKLVGVKVLGADSGSVSTIIGVDWV---VONKDYTG 259
 QY 341 VDVSYSVSGFTGTLVGEKYMQAIRALKAGIPMVVATGNVATSASSSMDLVANNHLKM 400
 Db 260 IRVINLSLGSSOSSDGTDSLQAVNNMADAGIYVCVAAGNSGPTMYTIGSPAAS---KV 316
 QY 401 TDTGNVTRTAHEDALIVASAKNOT-----VEBDKYNIGSESEFYKNIAGFF 447
 Db 317 ITVGAVD---SNDNIASFSSRGPTADGRLEPEYVAPGVDTIAPRASGTSNG--TPINDY 371
 QY 448 DK-----SKITTNEDGTA-----PSKLFYVYIGKODDOLIGDLGKTAIVMDRI 493
 Db 372 TKASGTSMAHPHVSQVGLLIQAHPSWTPDKVKTALIEA-----DIYAPKEIADIA 423
 QY 494 YTKDLKNAFKKAMDGARAIVVNTVNYNRDNTLPAWGEDEGTSQVFSISGDDG 553
 Db 424 Y-----GAGRVVYKAIRY---DVAKLFTFGSVADKSGATHTFPVSGATF 466
 QY 554 VK---LMMNIPDKKTEYKRNKKEDEKDLFOYYPIDMESNSNKPVNGDEKEIDFKFAP 610
 Db 467 VTATLYMDTGSSIDLXYLDXNGNEVDYSYRAYYGFEKVGYNPTAGTWIKVVSYGAA 526
 QY 611 DTDKEIKEDIIIVAGSTS 629
 Db 527 N-----YQVD-VVSDGSL 539

Search completed: December 7, 2001, 11:46:24
 Job time: 59 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 11:45:25; Search time 19.37 Seconds

(without alignments)
2611.249 Million cell updates/sec

Title: US-09-590-991-6

3454

Sequence: 1 YPVVLADTSSESDALNISDK.....PGKNIKSTLVINGKSTGY 664

Scoring table:

BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

PIR_68:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	453	13.1	1946	2 JCG6032	lactocepin (EC 3.4
2	434.5	12.6	1962	2 JCG6032	lactocepin (EC 3.4
3	429.5	12.4	1902	1 B44858	lactocepin (EC 3.4
4	427.5	12.4	1902	2 S06997	lactocepin (EC 3.4
5	422.5	12.2	1902	2 B45764	lactocepin (EC 3.4
6	343.5	9.9	806	2 A41341	microbial serine p
7	298.5	8.6	757	2 C84120	subtilisin-type pr
8	293	8.5	1167	1 A35066	streptococcal C5a
9	285.5	8.3	799	2 G83753	subtilisin-type pr
10	256	7.4	731	1 A55800	subtilisin-type pr
11	250	7.2	801	2 T07617	protease TMP - L
12	248	7.2	404	2 A44833	lactocepin (EC 3.4
13	242	7.0	754	2 T06577	subtilisin-like pr
14	239.5	6.9	1433	1 A36734	subtilisin-like pr
15	238.5	6.9	745	2 JCG6119	subtilisin-like pr
16	234	6.8	736	3 JCG518	subtilisin-like pr
17	232	6.7	856	2 T10585	serine protease
18	231.5	6.7	816	2 T06878	serine protease
19	228	6.6	747	2 T06580	subtilisin-like pr
20	227	6.6	754	2 T06579	subtilisin-like pr
21	219.5	6.4	766	2 T48389	cucumisin-like pro
22	219.5	6.4	815	2 T00538	probable serine pr
23	211.5	6.1	761	2 T07169	subtilisin-like pr
24	208	6.0	783	2 T01015	probable subtilisi
25	206	6.0	779	2 C86335	hypothetical prote
26	202	5.8	745	2 T07184	subtilisin-like pr
27	201.5	5.8	894	2 F69730	cell wall-associat
28	196	5.7	706	2 T01351	subtilisin-like pr
29	194	5.6	635	1 A29358	cerevisin (EC 3.4.

30	191	5.5	401	2 I39974	serine proteinase
31	188.5	5.5	381	2 JH0778	subtilisin (EC 3.4
32	187	5.4	645	1 SUBSMP	serine proteinase
33	186	5.4	1905	2 T18267	multidrug resistin
34	185.5	5.4	381	1 SUBS	subtilisin (EC 3.4
35	185.5	5.4	381	2 JQ1487	subtilisin (EC 3.4
36	184	5.3	769	2 D86335	subtilisin (EC 3.4
37	181.5	5.3	381	1 SUBS1	hypothetical prote
38	181.5	5.3	832	1 C86431	subtilisin (EC 3.4
39	180.5	5.2	1332	2 S41552	hypothetical prote
40	179	5.2	2510	2 T28160	probable transcrip
41	178.5	5.2	1398	2 T28159	hypothetical prote
42	177.5	5.1	764	2 T05768	pyrolysin (EC 3.4.
43	176	5.1	739	2 T12964	subtilisin-like pr
44	176	5.1	779	2 T07170	subtilisin homolog
45	175.5	5.1	319	2 I39866	subtilisin-like pr
					microbial serine p

ALIGNMENTS

RESULT	1	
JCG6032		
lactocepin (EC 3.4.21.96) precursor [similarity] - Lactobacillus delbrueckii subsp. b		
N:Alternate names: cell envelope-associated serine proteinase prtp		
C:Species: Lactobacillus delbrueckii subsp. bulgaricus		
C>Date: 31-Dec-1996 #sequence delbrueckii subsp. bulgaricus		
C:Accession: JCG6032		
R:Gilbert, C.; Atlán, D.; Blanc, B.; Portaller, R.; Germond, J.E.; Lapierre, L.; Molli		
J. Bacteriol. 178, 3059-3065, 1996		
A:Title: A new cell surface proteinase: Sequencing and analysis of the prtb gene from		
A:Reference number: JCG6032; MUID:96236017		
A:Accession: JCG6032		
A:Molecule type: DNA		
A:Residues: 1-1946 <GIL>		
A:Cross-references: GB:U48487; MID:96013471; PID:AA041529.1; PID:91381114		
A:Experimental source: NC001489		
A:Note: neither the complete nucleic acid sequence nor the complete translation are s		
C:Genetics:		
A:Gene: prtb		
C:Superfamily: lactocepin; subtilisin homology		
C:Keywords: hydrolase; serine proteinase		
F:1-34/Domain: signal sequence #status predicted <SIG>		
F:35-1946/Product: cell surface proteinase #status predicted <MAT>		
F:213-631/Domain: subtilisin homology #status atypical <SBT>		
Query Match	13.1%;	Score 453; DB 2; Length 1946;
Best Local Similarity	25.7%;	Pred. No. 2.7e-15;
Matches 175; Conservative 114; Mismatches 261; Indels 130; Gaps 28;		
OY	5	LADTSSESDALNISDKKVAENKEKHHSAMETSDPFEKKT-----AVIKEEYVSK 59
DB	27	LACTFQSAFAFVQAASQDTPSPRSASRAALTKYLDQEQRYKVAKSYKFDPAARQRAQSG 86
OY	60	NPIYDNTSNEAKIKKEENSKSGC-DYTDSPVKNKNTENPKKEDKVVYIAEFDKESGER 118
DB	87	QAVSKKNSSVRIVY---SLNKAAPDHT-----SKPTGSAASAKKIKQASD-QYKQOGR 138
OY	119	AIKELSLKNTKVLVYYDRIENGSALETTPDNLDKIKQIGISSVERAOKVOPMMNHARK 178
DB	139	VIKQVEEITGNKVRQFGYLVNAFISIDMDDDIDKVDLPQVKNVPPVKVYHP----- 191
OY	179	EIGVEEADIDKLSINAPF-GKNFDGKRGVNSINDTGTDRHKMKARIDDD-----AKASMR 232
DB	192	---TDESADQMAQVQDVQWQCKLKGEGMVISITDTGIDSSHQDKLDSGVSTALSSEVE 248
OY	233	FKKEDLKGTDKNKLWSDKIPHAFFYNGKITVEKYDDGHRDYPDPMHITAGLAGNDTE 292
DB	249	SDSKSL-GHGKRY--TEKYVGYNVADKNDQIV---DNCGC--EMHGHVAGTAGAN--- 297
OY	293	QDIKKNFGIDGICAPMAQIESYKMS--DAGSGFAGDETFPAIEDSTKHNVDVSVSSGFT 351

298 -----GOVGVAPDAQLAMKVFNSNMAKNSGAYDDDIISAIEDSVKLGADIVNNSLSGV 351
OY 352 GTGL-VGEKYMQAIRAKRAGIPVAVATGNVATSSASSSSMDLVAN---NHLKMTDQVNT 407
DB 353 SSDVSPDPQOQAAVAKASBAGIYINISAGNSGVAAGSTADGNVYNTGSELSTVTPGVY 411
OY 408 RTAAHEDAIAAVSAKNOQVEFDK--NIGESFKYRN--IGAEPDKSKITTNEDGTAKA-- 461
DB 412 -----PDALTVAASNSKVTDTYKDELGVTFSSNSSELKGA---AQVTTQLESNYSVL 462
OY 463 PSKLFVYVIGKQDDDLG---LDLRGKIAVMDR--ITTKOLKNARKKAKMDKARAIIMVYN 517
DB 463 TKRLVLVMDGLGADDTAAEKKAAYKGGOLAAVYKRGATYSAAVANAKA--AGAAAGIYIN 520
OY 518 TVNYYNRDMWTELPAWGAEADGCTKQVFSISGDGCVKLMMNINPDCKTEYVRNNKDEPK 577
DB 521 -----SEDG--LLSLSLDDKTFPTLGMKADGK 547
OY 578 DKLEQYPIDMESFNSKNPNVGE---KEIDFKFAPDPTDKELYKEDIIVPAGSTSMGPRI 634
DB 548 FKLKQOKKVRASRLKFGTALIDNSRAKMSDE-----TSMGPPTP 586
OY 635 DLLKRPDVSABGNKNTSLN 654
DB 587 ELDFKPEITAPGCKITYSLAN 606

RESULT 2

A32634
lactococin (EC 3.4.21.96) precursor [similarity] - Lactococcus lactis subsp. cremoris (S
N;Alternate names: cell envelope-associated serine proteinase prtP;
C:Species: Lactococcus lactis subsp. cremoris
C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 21-Jul-2000
C:Accession: A32634
R:Yos, P.; Simons, G.; Slezep, R.J.; de Vos, W.M.
J. Biol. Chem. 264, 13579-13585, 1989
A:Title: Primary structure and organization of the gene for a procaryotic, cell envelope
A:Reference number: A32634; MUID:89340435
A:Accession: A32634
A:Molecule type: DNA
A:Residues: 1-1962 <YOS>
A:Cross-references: GB:J04962; GB:M26310; NID:g149476; PIDN:AAA03533.1; PID:g149478
A:Note: The authors translated the codon GGT for residue 1103 as Tyr
A:Note: Part of this sequence, including the amino end of the mature protein, were deter
A:Genetics:
C:Superfamily: Lactococin; subtilisin homology
C:Keywords: hydrolase; serine proteinase; transmembrane protein
F:1-33/Domain: signal sequence #status predicted <SIG>
F:34-187/Domain: propeptide #status predicted <PRO>
F:188-634/Product: serine proteinase, cell-envelope-associated #status predicted <MAT>
F:1938-1955/Domain: subtilisin homology #status atypical <SBT>
F:1938-1955/Domain: transmembrane #status predicted <TM>

Query Match 12.6%; Score 434.5; DB 2; Length 1962;
Best Local Similarity 27.0%; Pred. No. 2.4e-14;
Matches 164; Conservative 92; Mismatches 240; Indels 111; Gaps 25;

OY 78 NSNKSQGYTSPFNKNTENPKKEDKV---YIAEFKKEGSEKAIKLSLKNTKLYLT 134
DB 84 NKVQODDIYVIVYOMSA-PASENGILRTDYSSFAEIODETNNVIAOAQSVKAVEQVT 142
OY 135 -----YDRIFNGSAITTPDNLDKIKOEGISSVEAEAKVOPMNHARKREIGEAID 187
DB 143 QOTAGESTGYVNGESTVRYVVDIPKLOAGVKTIVTLAKYVTPPDAAKANSMAVQAQVWS 202
OY 188 YLKSTINAFGKNGFGRGVANISIDTGDYRHKAMRIDDAKASRFFKEDL-KGTD--K 243
DB 203 NYK-----YKGEGVIVSVIDSGIDPTHKRLSD--KDKVLTGKSDVEKFTDVKH 251
OY 244 NYKLSDKTPAFNRY-NGKFTVEKYDDGRDYFDPHGHHIGIILAGNDEQIDKNFNSCID 302

DB 252 GRFNSKVPYGFENYADNNDTTDDKVD-----EQHGMHVAIGIANGCTGDD--PAKSV 303
OY 303 GIAPPAQIESYKMYSDAG-SGFAGDETMFHAIEDSIKHNVDVSVSSGE-TGTGLYGEKY 360
DB 304 GVAPEAQLAMKVFNSNSTSAKTGSATVSAIEDSAKIGADVLNLSLNSGNQTLDEPE 363
OY 361 WQAIARLKRAGIPVAVATGNVATSSASSSSMDLVANNHLKMTDQVNTPTAAHEDAIAAVS 420
DB 364 LAAVGNANESGAAVIAISAGNSTGSSAT--EGVNNDDYGLQDNEVSGPISRGATVYAS 421
OY 421 AKNOTVEFDKNI-----GGESFKY--RNIGAFDPKSKITTNEDGTAKPSKLEFYI 470
DB 422 AENDTVIQAAYITDGTGLQGLPEPTIQLSHDFTGSPFOKRYIYKADSG-----NL 473
OY 471 GKGDQDGLIGDLRGKIAVMDR--ITTKOLKNARKKAKMDKGAIRAIIVNTVYNRDMTE 529
DB 474 SKGALADYTA-DAKGIAYVRGSESPDKOKYAA--AGAAGLIYVT-----DGT 523
OY 530 LPAMGAEADGCTKQVFSISGDGCVKL--WMNINPDCKTEYVRNNKEDFKLEQYPID 587
DB 524 TPTMSIAL--TTPFTGLSSVTGOKLVDWYTAHPDSDLGK----- 563
OY 588 MESENSKNPNVGEDEKEIDFKFAPDPTDKELYKEDIIVPAGSTSMGPRIIDLKPDVSAF 647
DB 564 -----ITLMLPN---OKYTEDKM--SDFTSYGVSMLSEKPPITAPG 602
OY 648 NIKSTLN 654
DB 603 NIMSTON 609

RESULT 3

B44858
lactococin (EC 3.4.21.96) precursor [validated] - Lactobacillus paracasei (strain NCT
N;Alternate names: cell-envelope-associated proteinase prtP; serine proteinase Lp151
C:Species: Lactobacillus paracasei
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-May-2000
C:Accession: B44858; C44858; A44850
R:Holck, A.; Naes, H.
J. Gen. Microbiol. 138, 1353-1364, 1992
A:Title: Cloning, sequencing and expression of the gene encoding the cell-envelope-a
A:Reference number: A44858; MUID:92381481
A:Accession: B44858
A:Molecule type: DNA
A:Residues: 1-1902 <HOL>
A:Cross-references: GB:M83946; NID:g149580; PIDN:AAA25248.1; PID:g149582
A:Note: sequence extracted from NCBI backbone (NCBI:112261, NCBI:112263)
A:Note: the source is designated as Lactobacillus paracasei subsp. paracasei
A:Accession: C44858
A:Molecule type: protein
A:Residues: 'X',189-196 <HOL2>
R:Naes, H.; Nissen-Meyer, J.
J. Gen. Microbiol. 138, 313-318, 1992
A:Title: Purification and N-terminal amino acid sequence determination of the cell-w
A:Reference number: A44850; MUID:92226694
A:Accession: A44850
A:Status: preliminary
A:Molecule type: protein
A:Residues: 189-196 <NAE>
A:Cross-references: PIDN:AA22052.1; PID:g248666
A:Experimental source: strain NCDO 151
A:Note: sequence extracted from NCBI backbone (NCBI:94706)
C:Genetics:
C:Superfamily: Lactococin; subtilisin homology
C:Keywords: hydrolase; serine proteinase; transmembrane protein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:23-187/Domain: propeptide #status predicted <PRO>
F:188-1902/Product: serine proteinase, cell-envelope-associated #status experimental
F:208-634/Domain: subtilisin homology #status atypical <SBT>

Query Match 12.4%; Score 429.5; DB 1; Length 1902;

C:Genetics:
 A:Genome: plasmid
 C:Superfamily: lactocepin; subtilisin homology
 C:Keywords: extracellular protein; hydrolase; membrane bound; serine protease
 F:1-33/Domain: signal sequence #status predicted <SIG>
 F:34-187/Domain: propeptide #status predicted <PRO>
 F:188-1902/Product: serine protease, cell-envelope-associated #status experimental <M>
 F:208-634/Domain: subtilisin homology #status atypical <SBT>

Query Match 12.2% Score 422.5; DB 2; Length 1902;
 Best Local Similarity 27.0%; Pred. No. 9,6e-14;
 Matches 164; Conservative 91; Mismatches 241; Indels 111; Gaps 26;

78 NSNKGQDYTSFVNKNENPKED--KYYIAEFKKEGEGAKIKLSLKTKVLT 134
 DB NKVQDDIYVIVQMSAA-PASENGTLRTDYSSAEIQETNKYIAQAASVKAQEV 142
 135 -----YDRFNGSAIETTPDNDIKIKQEGISVERAKVQPMNHAREIGVEA 187
 DB OOTAGESGYVNGSTFRVVDIPKIKQAGVKTVTLAKYYPTDAKANSMAVQAV 202
 143 YLKSINAPFGKNPDGRGVISNIDTGDYRHKAMRIDDAKASMKREKED--K 243
 DB YLKSINAPFGKNPDGRGVISNIDTGDYRHKAMRIDDAKASMKREKED--K 243
 203 NYK-----YKGEGVVSVIDSGIDPRTKMDRLSD--KDYKLTGSDVEKEFTDAKH 251
 DB NYWLSDKIPHFANY-NGCKITVEKYDGRDYPDHGMHAGIAGNTEEDIKNFND 302
 244 GRYSNKPFGFNTADNDTIT---DQTVV--EDGHNVAGIIGANTGD--PAKSV 303
 252 GIAPNAOIFSYKMYSDAG-SGFAGDETFHAIEDSIKHNVDVSVSGF-TGTGLVGEKY 360
 DB GIAPNAOIFSYKMYSDAG-SGFAGDETFHAIEDSIKHNVDVSVSGF-TGTGLVGEKY 360
 304 GVAPDAQLLAKVFNTSDTSATGSSSTLSAIEDSAKIGADVLMNSGSDGNQTLDEPE 363
 DB GVAPDAQLLAKVFNTSDTSATGSSSTLSAIEDSAKIGADVLMNSGSDGNQTLDEPE 363
 361 WQAIRALKKAIPIWVATGATATSSSSMDLVANNHKKMTDGTGNTVTTAAHEDAI 420
 DB WQAIRALKKAIPIWVATGATATSSSSMDLVANNHKKMTDGTGNTVTTAAHEDAI 420
 364 LAAVQANESGTAIVISAGNSGTSAT--EGVKNDDYGLDNDNEVGTPTGSRGATTVAS 421
 DB LAAVQANESGTAIVISAGNSGTSAT--EGVKNDDYGLDNDNEVGTPTGSRGATTVAS 421
 421 AKNOVERDKYNI-----GGESEFY--RNIGAFDPKSKITTNEDTKAPSKLFY 470
 DB AKNOVERDKYNI-----GGESEFY--RNIGAFDPKSKITTNEDTKAPSKLFY 470
 422 AENDVITQAVTITDGTQLQGPETIQSLSSNDPFGSPQKFFVYVKDSG-----NL 473
 DB AENDVITQAVTITDGTQLQGPETIQSLSSNDPFGSPQKFFVYVKDSG-----NL 473
 471 GKGGDQDLIGDLGKIAVMDR-IYTRDLKNAFKKAMDGAIRAIWVTVYNNRDMTE 529
 DB GKGGDQDLIGDLGKIAVMDR-IYTRDLKNAFKKAMDGAIRAIWVTVYNNRDMTE 529
 474 SKGALADYTA-DAKGKIAIVKRGELSFDDKQKYAQA--AGAAGLIIV-----NDG 525
 DB SKGALADYTA-DAKGKIAIVKRGELSFDDKQKYAQA--AGAAGLIIV-----NDG 525
 530 LPANGYDEGTSQGVVISGDDGVKL--WMMINPDKTEYKRNKKEDEFKLEQYFID 587
 DB LPANGYDEGTSQGVVISGDDGVKL--WMMINPDKTEYKRNKKEDEFKLEQYFID 587
 526 VTSMAIT---TTPPTGELSVTGOKLVDVTAHPDDSLGK----- 563
 DB VTSMAIT---TTPPTGELSVTGOKLVDVTAHPDDSLGK----- 563
 588 MESFNSNKNPVGDEKEIDFKAPDDELYKEDIIVPAQSTSMGPRIDLLKPDVSA 647
 DB MESFNSNKNPVGDEKEIDFKAPDDELYKEDIIVPAQSTSMGPRIDLLKPDVSA 647
 564 -----TALTVLPN---QKYTEKMA--SDFTSYGVSNLSKFKPITABGG 602
 DB -----TALTVLPN---QKYTEKMA--SDFTSYGVSNLSKFKPITABGG 602
 648 NIKSTLN 654
 DB NIKSTLN 609
 603 NIMSTON 609

RESULT 6

Microbial serine protease (EC 3.4.21.-), minor (vpr), precursor - Bacillus subtilis

C:Species: Bacillus subtilis
 C:Date: 17-Jul-1992 #sequence, revision 17-Jul-1992 #text, change 20-Jun-2000

C:Accession: A41341; B41341; S59700; D69730
 R:Stroma, A.; Rulo Jr., G.A.; Theriault, K.A.; Dwyer, M.; Wilson, S.W.; Pero, J.

J. Bacteriol. 173, 6889-6895, 1991
 A:Title: Cloning and characterization of the gene for an additional extracellular serine

A:Reference number: A41341; MUID:92041574
 A:Accession: A41341

A:Molecule type: DNA
 A:Residues: 1-806 <SLO>

A:Cross-references: GB:M76590; NID:g143819; PIDN:AAA22881.1; PID:g143820
 A:Accession: B41341
 A:Molecule type: Protein
 A:Residues: 161-195 <SL2>
 R:Glaser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales, W.; Hujlo, M.F.; Ionesi, A.; Rapoport, G.; Danchin, A.
 Mol. Microbiol. 10, 371-384, 1993
 A:Title: Bacillus subtilis genome project: cloning and sequencing of the 97 kb region
 A:Reference number: S39655; MUID:95020537

A:Accession: S39700
 A:Molecule type: DNA

A:Residues: 1-806 <GLA>
 A:Cross-references: EMBL:X7124; NID:9413923; PIDN:CAA51601.1; PID:9580871

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Beck, C.; Broer, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fultz, C.; Fujita, M.; Fuma, S.; Galizzi, A.; Galizzi, A.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hujlo, M.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardine, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portek, Rieger, M.; Rivolta, C.; Roeha, E.; Roeha, B.; Rose, M.; Sadale, Y.; Sato, T.; Scan, A:Authors: Schleicher, S.; Schroeter, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; St

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiy, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
 A:Reference number: A69580; MUID:98044033

A:Accession: D69730
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-806 <RUN>

A:Cross-references: GB:299123; GB:AL009126; NID:g2636240; PIDN:CAB15835.1; PID:g26365
 A:Experimental source: strain 168

C:Comment: The amino terminal sequence of the mature protein and a molecular weight c
 C:Genetics:
 A:Gene: vpr
 A:Start codon: TTG

C:Superfamily: microbial serine protease vpr; subtilisin homology
 C:Keywords: hydrolase; serine protease

F:1-28/Domain: signal sequence #status predicted <SIG>
 F:29-160/Domain: propeptide #status predicted <PRO>
 F:180-548/Domain: subtilisin homology #status atypical <SBT>

Query Match 9.9% Score 343.5; DB 2; Length 806;
 Best Local Similarity 23.4%; Pred. No. 3.1e-10;
 Matches 159; Conservative 89; Mismatches 200; Indels 231; Gaps 30;

35 SAMEISODFE-----KRTAVI---KEEVSKNPVIDNNTSNEAKIKEENS 79
 DB 31 ASSKTSADLEAEVFGDIDMTSKTYIVELKESLA-----EAKESAGE 77
 80 NKSQGDYTSFVNKNENPKEDKRVYIAEFKKEGEGAKIKLSLKTKVLTYYDRIF 139
 DB 78 SKSK-----LKTAPRKAK---NKAIR---AVKNGKVNREYQVE 110
 140 NGSALIEETPDNDIKIKQEGISSVE-----RAQKQPMNHAREIGV 182
 DB 111 SGFSMKLPANEIPKLLAVKDKAVYPNVYKTKDNMKKQDVITSEDAVSPQDDAPYIGA 170
 183 EADYDLYKSIAPFGKNPDGRGVISNIDTGDYRHKAMRIDDAKASMKREKEDLGT 242
 DB 171 NDANDL-----GYTGKIVAIITDGVENH-----PDLK--- 200
 243 KNYWLSDKIPHFANYNGCKITVEKYD-----DGDYEDPHGMHAGIAGNTEEDI 295
 DB 201 KN-----FGQYGYDVEDNDYDPKPTGDDPCEATDGTGVAAGVANGT--- 246
 296 KNFGGIGIAPNAOIFSYKMYSDAGSGFAGDETFHAIEDSIKHNVDVSVSGFTGTGL 355
 DB 247 -----IKGVADATLLAYRYLGPGGSGTT--ENVIAGVERAVVDGADVMMNLISG-----NS 295

```

0Y 356 VGEKWMQIIRLRK---GIPWVAATGNATSASSMDIVANNHLKMTDGTANTPTAAH 412
Dh 296 LNNPDMATSTLDDMAASGCVAAVTSNGN---SCPNGH-----TYGSPCTS 337
0Y 413 EDATAVASAKNQVEFEDKVNIGESFKYRNIGAFEDSKIT--TINEDGTNA--PSKLKFV 468
Dh 338 REALSVGATQLPUNNY -AVTFGS-----YSSAKVGVYKKEDVKALNNKEVELY 385
0Y 469 YTGQODDGLIGLIRGKIAVMDR--IYTKDLKNAFKKAMDKGARAIWVNTVYVYNDN 526
Dh 386 EAGIGEAKDEFGKDLTGKVAAYVKRGSIAFYVDKADAKA---GAIGWVYVNNLSGEIEAN 442
0Y 527 WTELPAMKYEADDEGCKSQSVFSISGDDGVKLLMMNINPDKKTVEYKRNKKEDEFDKLEQYPI 586
Dh 443 ---VPGMSV-----PTIKSLDEDEKL---VSALAGEETKTTFKLTVSKAL----- 482
0Y 587 DMESNSNKKPWNVGDEKELDFKFAPTDCKELKEDIIVASGTSWGPRI DLKPDVASP 645
Dh 483 -----GQVADP-----SSRGVPVMDTWIMKIPDISAP 508
0Y 646 GKNIKSTLNVINGKSTYGY 664
Dh 509 GVNIVSTIPTHDPHPHYGY 527

```

```

RESULT 7
CB4120
subtilisin-type proteinase (EC 3.4.21.-) BH3763 precursor [similarity] - Bacillus halodurans
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 19-Jan-2001
C:Accession: CB4120
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20263314
A:Accession: CB4120
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-757 <STO>
A:Cross-references: GB:AP001519; GB:BA000004; NID:G10176109; PIDN:BAB07482.1; GSPDB:GN00
C:Genetics:
A:Gene: BH3763
C:Superfamily: microbial serine proteinase vpr; subtilisin homology
C:Keywords: hydrolase; serine proteinase
P:1-20/domain: signal sequence #status predicted <SIG>

```

Query Match	8.6%	Score 298.5	DB 2:	Length 757
Best Local Similarity	21.6%	Pred. No. 5	8e-08	
Matches 127	Conservative 97	Mismatches 181	Indels 183	Gaps
OY	96 ENPK---	KEDKVVYIAEFKDEKESGEKAIKELSSLIKNTKLYTYDRIFNQSALETTPDND	1522	
Db	30 EKPRLADEDEIVYIVLTNRDLDTETIKDYER--	SLPSELSTMTFRLNGNSFOLPREBEVE	86	
OY	153 KIKQEGISSYERAKQKQOPMNHARKETGVEALDYKTSINAPKGNFQGRGWSINDP	212		
Db	87 KLTIVSYGERDHWITYEATINESAIPITGAQ----	VRGLMDEGVALITKQKRVAVIOT	142	
OY	213 GTDVRHKMRIRIDDDAKKSMRKEDLKGTDKNTKWLSDKIPHAENYNGKITYEKKIDCGR	272		
Db	143 GIDYTH-----	PDLOSSYKGGVDFVD--VDD--	166	
OY	273 DYFDP-----	HGMHIAGITLGNPDTEODIKNFNIDGIAPAQIYSKMYSDAG	320	
Db	167 ---DPMETIASQGPPTLHGTHTVSIIAN-----	GGVKCAVPAPELTYAVRALGPGC	214	
OY	321 SGFAGDEIMFAIEDSIKHNADVSVSSGFLGTGLVEAKTQWQAIRLKA--	GIPMYVA	377	
Db	215 QGTT--EQQVIAIKAEVGDGVYNTLSIGTVN-----	PDMPISLADAIVBEQVAAYS	268	

[illegible]

```

RESULT      8
A35066
Streptococcal C5a peptidase (EC 3.4.21.-) precursor - Streptococcus pyogenes
C:Species: Streptococcus pyogenes
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 18-Jun-1999
C:Accession: A35066; S52539
R:Chen, C.C.; Cleary, P.P.
J. Biol. Chem. 265, 3161-3167, 1990
A:Title: Complete nucleotide sequence of the streptococcal C5a peptidase gene of Stre
A:Reference number: A35066; MUID:90153964
A:Accession: A35066
A:Molecule type: DNA
A:Residues: 1-1167 <CHE>
A:Cross-references: GB:J05229; NID:g153776; PIDN:AAA26960.1; PID:g552009
A:Note: the amino end of the mature protein was determined by protein sequencing
R:Katerov, V.; Schalen, C.; Tocolian, A.A.
Mol. Gen. Genet. 245, 78-85, 1994
A:Title: Sequencing of genes within the v1r regulon of Streptococcus pyogenes type M1
A:Reference number: S52535; MUID:95147851
A:Accession: S52539
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-75, 'V', '77-78, 'K', '80-83, 'PS' <NAT>
A:Cross-references: GB:S75411; NID:g914107; PIDN:AA33264.1; PID:g914112
A:Note: In Genbank entry S75411, release 106.0, the initiation codon TTG for residue
C:Genetics:
A:Gene: scpA
A:Start codon: TTG
C:Function:
A:Description: specifically cleaves a Lys-Asp peptide bond near the carboxyl end of h
C:Superfamily: streptococcal C5a peptidase; subtilisin homology
C:Keywords: cell wall; hydrolase; serine proteinase; tandem repeat; transmembrane pro
F:1-31/Domain: signal sequence #status predicted <SIG>
F:32-167/Product: streptococcal C5a peptidase #status predicted <NAT>
F:121-526/Domain: subtilisin homology #status atypical <SBT>
F:1029-1139/Domain: cell wall spanning #status predicted <CMS>
F:1034-1101/Region: glycine/proline-rich 17-residue repeats
F:1040-1157/Domain: transmembrane #status predicted <TM>
F:130,153,512/Active site: Asp, His, Ser #status predicted

```

[illegible]

```

0Y 206 VTSIDTGTDRYHNKAMRIDDOKAKSMRKKEDLKTDKNY-----WISDKTRPHEFNYYN 259
Db 125 VVA:::VV:::DDGDKNHEMRLLDTKARYQ--SKVEDLEKAKKEHGTITGEWVNDKAAVYHDSK 183
0Y 260 GCKITVEKYDDGRDYFDPHGMHIAOTLAGNDPEODIKNFNGIDGTAIPNAOJFSYKM---- 315
Db 184 DOKTAVDQ-----EHGTHVSGILSN--APSEKTEPRLLEGAMPENQOLLMHVEIYN 233
0Y 316 --YSDAQSGFAGDEMFMAHIEDOSIKHNVDVSYSSSEFTGTGLVG--EKYQOALRALRKGI 372
Db 234 GLADYARNYA-----QAIRDAVNLGARYINNSFGNALAALAYNLPDETKKADYAKSRGV 287
0Y 373 PMVVTGNYATVSASSSSNDLVANNHKLMTDGTGNVTRTAHEDAIYA--SAKNQTEVERDK 430
Db 288 SIVTSAGNDSRSGCKTRLPLA--DH---PDGVGVOTPPAADSTLTVASVSPKQJLTERAM 342
0Y 431 VNIGGESKRYRNIGAFPFDSKTTTNETDGTAKPSK-LKTVYVIGKGODDILGLDLNGKTAV 489
Db 343 VKTDDOODK-----EMPYLSTNR--FEBPNKAYDYAANRKGKEDDF-KDVGKITAL 390
0Y 490 MORITVTKDLKNAFKAMDKGARAIWVNTVNYNYND-----WMTLEPAMGYEADSGT 541
Db 391 IER-GDIDFKCKVNAAKKAGAVGLI-----YDNODKGFPIELPVDQMPAA----- 436
0Y 542 KSQVFSIGDDGVKILMNIINPDKTEVRKNNKEDFKLEQYYPIDMESFNKNPVGDE 601
Db 437 -----FISRKGLLLKLD--NP-----Q 451
0Y 602 KEIDFKFAPDFTKELYKEDIIVPACS-----TSMGPRIDLLKPDVSAPKNIKISTLN 654
Db 452 KTIITFNAPRK-----VLPFASGJKLSRFSMSGLTADONIKPDIAPBOQDILS--S 499
0Y 655 VINGK 659
Db 500 VANNK 504

```

RESULT 9
GB3753
subtilisin-type proteinase (EC 3.4.21.-) vpr precursor [similarity] - *Bacillus halodurans*
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 19-Jan-2001
C:Accession: GB3753
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirose, T.
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: AB3650; MUID:20265314
A:Accession: GB3753
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-799 <STC>
A:Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA04550.1; GSPDB:GN000000000
C:Experimental source: strain C-125
C:Genetics:
A:Gene: vpr
C:Superfamily: microbial serine proteinase vpr; subtilisin homology
C:Keywords: hydrolase; serine proteinase
;1-29/domain; signal sequence #status predicted <SIG>

Query Match	8.38	Score 285.5	DB 2	Length 799
Best Local Similarity	21.48	Pred. No. 2.9e-07		
Matches 140; Conservative	96	Mismatches 196	Indels 221	Gaps 30

```

0y 63 IDNNNSNEAKIKE-ENSN-----KSGDGYDPSFVNKNTENPKREDKVYIAEFKDRSGE 117
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::
Db 43 IDTSSSALEYIVELIEDPSITAEKHQG-----OKOSNELKAKAROSVLEIQD----- 90
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::
0y 118 KAKELSLKNTKVLVYTDRIENGSALETTPNDLKITQIEGISSVERBAQKVCPMMNHAR 177
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::
Db 91 -----LVPSSTVTHEDFDLFGSALEPAHQIPSLIGIDGVAV--YPIIEVEVTFD 141b
    ||::||::||::||::||::||::||::||::||::||::||::||::||::||::||::

```

```

QY 178 KEIWE-----EADIDYKLSINPF-GKN-----PDGBMYISNIDTDYDRHKARID 224
Db 142 DEVEIEKDAYSPMLD-----SAPFTGANAMEAGTGBEITVAIIDTDGVYDH-----190
QY 225 DDAKASMRKKEDLJKCTDKNKYWLSDRI PHAFNYYNGSKITVEKYDGRDY -FDP-----277
Db 191 -----PDLVHAFGDKYKGMDF-IDNNDDPQETPPCGPRGIET 225
QY 278 -HGMIHAGLACNDTQDIDIKNENGIDGIAPNAQIFSKYKMSDAGSGAGDETHALIEDS 336
Db 226 TIGTHVAGVVAANGL-----IKGVAPDADULAIRVLGPGGRGTAG--VYAGIERA 274
QY 337 IKHNVDYVSVSSGFTGTGVLGEKRYWQAIRAKRAGIPMYATGNYATSASSSSWDLVANN 396
Db 275 VODGADIMLHLSIGNT-----LNDPWFATSI-ALDW-AMAEG 308
QY 397 HLKMTDIGNVTRIAHEDAIIVASAKNQT-----EFDKNIGESKKYRNIGAFPRKSK 451
Db 309 VVAVYISNGN-----SGPNMYTVGSPGTSRDALISVATRLPYNKY-----KAS 350
QY 452 IYTNEDGTKAPS-----KLKFVYHGKQODDILGIDLRGKIAVMDRI 493
Db 351 VETS-DCIDIDYPSADIMGFPSDEBELLFEDGETYEFARGLKPDGFBEVDVBGKALLVRG 409
QY 494 YTKDLKNAFKKAMDKCARALIMVNYNYNYNRDMWTELPMANGYREAD-EGTYSQVFSISGD 552
Db 410 EIRPEEKA-----ENAKAAGAVGAIYYN-----VAGVOPVYVGLAIPITIMLANED 455
QY 553 GYKLMWMIIPDKKTEYKRRNNKEDFEDKLEQYYPIDMESFNNSKNPNVDEKEIDEKFPAPDT 612
Db 456 GIKMRHLELNGNT-----VTFISIEF 476
QY 613 DKELXKEDIIIVAGSTSMGPRIDL-LKRPVSAPGKNIKSTLWVINGKSTYYG 664
Db 477 DK-LVGEYI--ADFSRRGPVMTIMIKIPVSAAGVAIYSTIPTHODPDYGY 525

```

```

RESULT 10
A:55800
cucumislin (EC 3.4.21.25) precursor - muskmelon
C:Species: Cucumis melo (muskmelon)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A55800
R:Yanagata, H.; Masuzawa, T.; Nagaoaka, Y.; Ohnishi, T.; Iwasaki, T.
J. Biol. Chem. 269, 32725-32731, 1994
A>Title: Cucumislin, a serine protease from melon fruits, shares structural homology
A:Reference number: A55800; MUID:95105149
A:Accession: A55800
A:Molecule type: mRNA
A:Residues: 'LIAKSTT',1-731 <YAM>
A:Cross-References: EMBL:D32206
A:Note: It is uncertain whether Met-1 is the initiator or whether translation is initiated by Met-1.
C:Superfamily: subtilisin-like proteinase ag1; subtilisin homology
C:Keywords: hydrolase; serine proteinase
F:1-110/Domain: signal sequence and amino-terminal propeptide (fragment) #status predicted
F:111-615/Product: cucumislin #status predicted <AM>
F:111-533/Domain: subtilisin homology #status atypical <SBT>
F:615-731/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:615-616/Cleavage site: Thr-Gly (autolytic) #status predicted

```

Query Match	7.48;	Score 256;	DB 1;	Length 731;
Best Local Similarity	24.1%;	Pred. No. 8.4e-06;		
Matches 148;	Conservative 94;	Mismatches 211;	Indels 162;	Gaps 31;

```

Qy 94 NTEPNKDEKQVYIA-EKKDESG--EKAKEL---SLKNTKYLYTPYRFGSAIET 146
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Db 25 DSDDDGKIYIYVNGRKLEDDSDSLHNRALBEOVGSTPAESVLYHTYRSGFGPAVL 84
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
147 TPNDLRIKTKQIEGISSVERAOKVOPMMNHAKELGVEEAIDYLSINAPFGKNDGGMV 206
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
85 TEEEAKEKIASNEGVSU-----PLNE-INELHTTRSMDFL-----GPLYTPRPSQV 130
      ::::: ::::: ::::: ::::: ::::: ::::: ::::: :::::
Qy

```



```

Db 113 AVOANNEGTAIAVISAGNSGTSGSAT--EGVKNKYDYLQDNEMVCTPCTSGATTVASAE 170
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 423 NOTVEFDKVN-----GGESEFKY--RNTGAFEDKSKITTNEDGTAPSKLKFVYIGK 472
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 171 NTDVITTAQVTTTDTGTGLQGLPETIQLSSNDFTSGDOKFFVYVVDASGNLSK-----GK 224
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 473 GQDDDLGLDLGRKIAVMDRIYTKDLKNAFKAMDKGARAIMVY-----VTVANYNNDN 526
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 225 VADT---ADAKGKIATYKR---GELTFADKQ---KYAQAGTYPDLKKNKRTGQYIGGM 275
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 527 WTELPMAGYEADGKTSQVFSISGDDGVKLMNMIN 561
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 276 VTDA-----DANOIYDDQAIASFSSDKNA-LYNEIS 304
|: : : : : | : : : : : | : : : : : | : : : : : |

```

RESULT 13

```

T06577
subtilisin-like proteinase (EC 3.4.21.-) - tomato
N:Alternate names: proteinase p69c
C:Species: Lycopersicon esculentum (tomato)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
C:Accession: T06577
R:Melchior, J.; Amrhein, N.; Schaller, A.
submitted to the EMBL Data Library, May 1998
A:Description: The gene family of subtilisin-like proteases in tomato.
A:Reference number: 215771
A:Accession: T06577
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-754 <MEI>
A:Cross-references: EMBL:AJ005171; PIDN:CA06412.1
A:Experimental source: cultivar VFM8
C:Genetics:
A:Gene: p69c
A:Superfamily: subtilisin-like proteinase ag12; subtilisin homology
C:Keywords: hydrolase; serine proteinase

```

Query Match 7.0%; Score 242; DB 2; Length 754;

Best Local Similarity 22.58; Pred. No. 4.6e-05;

Matches 140; Conservative 92; Mismatches 236; Indels 154; Gaps 28;

```

Oy 83 QGDYDTSFVNKNNTENPKKEDKVYIAEFKDES-----GEKAIKELSLKNTKVLVT 134
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 21 QSDDETYIV--HVESP--ESLITQSSFMDELYSLFPEPMASISSGNEBASITYS 76
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 135 YDRIFNGSAITTPDNLKIKQIEGSISSVERAQVQPMNHARKEIGVEAIDYLSINA 194
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 77 YHNWMTGFAARLTAEQVYEMERKKGFSVAKQKRLTLHTTTPSFLIQKNGVWKDSNY 136
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 195 PFGKNPDSGRGVINIDPGTGYRHKAMRIDDAKASMKFKEDLKTGDKNW---LSDKI 251
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 137 -----GKGVIIGVLDGLIIPDHPSP-----SDVGMPSPPAKMGVCSSNFTKCNKL 184
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 252 PHAENYNGKITVEKYDDGDYEDPHGMHAGILAGNDEODIK-----NFGCID-GI 304
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 185 IGANSYELGNMSPIDN-----DGHGTHASTAG-----AFYKGAHVHGNANGTANGV 232
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 305 APNNOISYKAYS DAGSFGAG---DETWFHAIEDSIKHNVDVVSSEFTGTGLVEKYW 361
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 233 APLAHIAIKYK-----CGFDPKCPGSDILAMDAIDGVDILSISLGSLSPLYDETIA 287
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 362 QAIRALKRAGIPMYVATGNGVNTSASSSSWMDLVANNHLKMTGTGVTTRAHEDIAVNSA 421
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 288 LGAVSTORGLIVSCSAGNSGSPAS-----VDNSAPILTVGASTLDRKIKATVKKLNG 342
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 422 KNQTVEDPKVNGESF---KYRNIGAF--FDKSKITTNEDGTAPSKLKFVYIGKGDQ 476
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 343 E-----EFG--GSAVHPKTSNATFTLEDAK-----NADPSETPYCRGRGSLDP 387
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 477 DLIGLDLGRKIAV---MDRIYTKDLKNAFKAMDKGARAIMVYVTVNY--YNRDNWTEL 531
|: : : : : | : : : : : | : : : : : | : : : : : |

```

```

Db 388 -----AIFRKITVLCLAFGVANVDKGAVK---DAGVGMIIVINPSOYGVTKSADAHVLP 439
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 532 AMGEADGKTSQVFSISGDDGVKLMNMINP-----DKKTEVKNKNNEFDKX 579
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 440 ALVVSADGKIKIRATVNS-----ILNPAVITTFGCTIIGDKNAPI----- 479
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 580 LEQYYPIDMESEFNKPNVDEKEIDPFKAPDTKELYKEDIIVPAGSTSGPRIDLLK 639
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 480 -----VAAFSSRGPRP-----ASNGILKPOLI-----GPGVNIILAA 510
|: : : : : | : : : : : | : : : : : | : : : : : |
Oy 640 PDVSAFG-KNIKSTLVNIGKS 660
|: : : : : | : : : : : | : : : : : | : : : : : |
Db 511 WPTSVGKNKNTKSTFNISIGTS 532
|: : : : : | : : : : : | : : : : : | : : : : : |

```

RESULT 14

```

A36734
bactillopeptidase F (EC 3.4.21.-) precursor bpr [validated] - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 16-Jun-2000
C:Accession: A36734; A35131; A35750; S08223; JN0335; I39849; B69596; JU0084
R:Sloma, A.; Rufo Jr., G.A.; Rudolph, C.F.; Sullivan, B.J.; Theriault, K.A.; Pero, J.
Bacteriol. 172, 5520-5521, 1990
A:Reference number: A36734; MUID:90368623
A:Contents: erratum
A:Accession: A36734
A:Molecule type: DNA
A:Residues: 1-1433 <SILO>
A:Cross-references: GB:M29035; NID:g143307; PIDN:AAA62679.1; PID:g143308
R:Sloma, A.; Rufo Jr., G.A.; Rudolph, C.F.; Sullivan, B.J.; Theriault, K.A.; Pero, J.
Bacteriol. 172, 1470-1477, 1990
A:Title: Bactillopeptidase F of Bacillus subtilis: purification of the protein and c
A:Reference number: A35131; MUID:90170864
A:Accession: A35131
A:Molecule type: DNA
A:Residues: 1-365, 'S', 367-682, 'EIMP', 893, 'Q', 895-896 <SL2>
A:Cross-references: GB:M29035
A:Note: the authors translated the codon GAA for residue 545 as Leu
R:Wu, X.C.; Nathoo, S.; Pang, A.S.H.; Carne, T.; Wong, S.L.
J. Biol. Chem. 265, 6845-6850, 1990
A:Title: Cloning, genetic organization, and characterization of a structural gene en
A:Reference number: A35750; MUID:90216713
A:Accession: A35750
A:Molecule type: DNA
A:Residues: 1-392, 'V', 394-828, 'NIRRLYSLKFCRSRHSV' <MUA>
A:Cross-references: GB:J05400; NID:g142607; PIDN:AAA83363.1; PID:g1119197
A:Note: this sequence has been corrected
A:Accession: B35750
A:Molecule type: DNA
A:Residues: 876-935, 'CG' <MU2>
A:Cross-references: GB:J05400; NID:g142607; PIDN:AAA83363.1; PID:g1119197
A:Note: this sequence has been corrected
R:Masuda, E.S.; Anaguchi, H.; Sato, T.; Takeuchi, M.; Kobayashi, Y.
Nucleic Acids Res. 18, 657, 1990
A:Title: Nucleotide sequence of the sporulation gene spoIIIGA from Bacillus subtilis.
A:Reference number: S08223; MUID:90174995
A:Accession: S08223
A:Molecule type: DNA
A:Residues: 1410-1433 <MAS>
A:Cross-references: EMBL:X17344; NID:g40165; PIDN:CA35224.1; PID:g809661
R:Kato, T.; Yamagata, Y.; Arai, T.; Ichishima, E.
Biochim. Biotechnol. Biochem. 56, 1166-1168, 1992
A:Title: Purification of a new extracellular 90-kDa serine proteinase with isoelectri
A:Reference number: JN0335; MUID:93005071
A:Accession: JN0335
A:Molecule type: protein
A:Residues: 195-218, 'A' <KAT>
A:Note: source of this material was Bacillus subtilis (natto)
R:Beall, B.; Lowe, M.; Lutkenhaus, J.
J. Bacteriol. 170, 4855-4864, 1988
A:Title: Cloning and characterization of Bacillus subtilis homologs of Escherichia co
A:Reference number: I39849; MUID:89006108
A:Accession: I39849

```


Fri Dec 7 13:21:48 2001

us-09-590-991-6.rpr

Page 10

[illegible]

Search completed: December 7, 2001, 11:48:40
Job time: 195 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 11:46:26 ; Search time 13.44 Seconds
(without alignments)
1811.417 Million cell updates/sec

Title: US-09-590-991-6
Perfect score: 3454
Sequence: 1 YPVVLADTSSSEDALNISDK.....PGKNIKSTLVINGKSTYGY 664

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues
Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :: SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	434.5	12.6	1902	1 P3P_LACLC	P15592 lactococcus
2	429.5	12.4	1902	1 P2P_LACPA	Q02470 lactobacilli
3	427.5	12.4	1902	1 P2P_LACLC	P15293 lactococcus
4	421.5	12.2	1902	1 P1P_LACLC	P16271 lactococcus
5	343.5	9.9	806	1 SUBV_BACSU	P29141 bacillus su
6	293	8.5	1167	1 SCAL_STRPY	P15926 streptococ
7	292	8.5	1181	1 SCA2_STRPY	P58099 streptococ
8	239.5	6.9	1433	1 SUBF_BACSU	P16397 bacillus su
9	201.5	5.8	894	1 WPRB_BACSU	P54423 bacillus su
10	194	5.6	635	1 PRTB_YEAST	P09332 saccharomyc
11	191	5.5	401	1 THES_BACSP	Q45670 bacillus sp
12	188.5	5.5	381	1 SUBN_BACNA	P35833 bacillus su
13	187	5.4	645	1 SUBE_BACSU	P16396 bacillus su
14	186	5.4	1905	1 TAGB_DICDI	P54683 dictyostell
15	185.5	5.4	381	1 SUBT_BACSA	P00783 bacillus su
16	185.5	5.4	381	1 SUBT_BACST	P29142 bacillus su
17	181.5	5.3	381	1 SUBT_BACST	P04189 bacillus su
18	180.5	5.2	1332	1 SPT7_YEAST	P35177 saccharomyc
19	175.5	5.1	319	1 ISPL_BACSU	P11818 bacillus su
20	175	5.1	379	1 SUBT_BACLT	P00780 bacillus su
21	173	5.0	682	1 SUBT_BACAM	P07596 lactococcus
22	170	4.9	682	1 NISP_LACLA	P27693 bacillus al
23	168.5	4.9	380	1 ELVA_BACAO	P28842 bacillus sp
24	167	4.8	420	1 SUBT_BACSU	P41362 bacillus su
25	163.5	4.7	380	1 ELVA_BACCS	P29139 paenibacilli
26	160.5	4.6	326	1 ISP_PABPO	P16588 vibrio algi
27	159	4.6	534	1 PROA_VIBAL	P17953 enterococcu
28	159	4.6	1296	1 ASAI_ENTFA	P20724 bacillus sp
29	158	4.6	378	1 ELVA_BACSP	P43153 clostridium
30	157.5	4.6	1304	1 COLA_CLOPE	Q00798 plasmodium
31	156	4.5	2869	1 RBPL_PLAVB	P07518 bacillus pu
32	155.5	4.5	275	1 SUBT_BACPU	P29145 hydra atten
33	153	4.4	710	1 NECH_HYDAT	

34	152	4.4	361	1 ELVA_BACHD	P41363 bacillus ha
35	152	4.4	944	1 NUF1_YEAST	P32380 saccharomyc
36	152	4.4	960	1 YMX6_YEAST	Q04279 saccharomyc
37	151.5	4.4	1167	1 CAGA_HELPU	Q92111 helicobacte
38	151	4.4	533	1 PEPC_ASPNC	P33295 aspergillus
39	151	4.4	1658	1 YMG7_YEAST	Q03611 saccharomyc
40	150.5	4.4	802	1 NAB3_YEAST	P38996 saccharomyc
41	149	4.3	1053	1 SLPN_BACBR	P06546 bacillus br
42	146.5	4.2	1186	1 CAGA_HELPU	P55980 helicobacte
43	146	4.2	477	1 PRCA_ANASO	P23915 anabaena sp
44	146	4.2	578	1 LIPA_MYCPU	Q50274 mycoplasma
45	145.5	4.2	1375	1 GTFC_STRMO	P13470 streptococc

ALIGNMENTS

RESULT 1
P3P_LACLC
ID P3P_LACLC STANDARD; PRT; 1902 AA.
AC P15292;
DT 01-APR-1990 (rel. 14, Created)
DT 01-APR-1990 (rel. 14, Last sequence update)
DT 15-JUL-1998 (rel. 36, Last annotation update)
DE P11-TYPE PROTEINASE PRECURSOR (EC 3.4.21.96) (LACTOCCEPIN) (CELL WALL-
DE ASSOCIATED SERINE PROTEINASE).
GN PRTP.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OG Plasmid.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_TaxID=1359;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 188-197.
RC STRAIN-SK11;
RA MEDLINE=89340435; PubMed=2760036;
RX Vos P., Simons G., Slezan R.J., de Vos W.M.;
RT "Primary structure and organization of the gene for a procaryotic,
RT cell envelope-located serine proteinase.",
RL J. Biol. Chem. 264:13579-13585(1989)
CC -!- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE
CC GROWTH OF THE BACTERIA ON MILK.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: J04962; AAA03533.1; ALT_SEQ.
CC PIR: A32634; A32634.
CC HSSP: P00782; 2SBR.
CC MEROPS: S08_019; -.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003137; PA.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF02225; PA; 1.
DR Pfam: PF00082; Peptidase_S8; 3.
DR PRINTS: PR00723; SUBTILASIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
KW Hydroxylase, Serine protease, Cell wall, Zymogen, signal, plasmid,
KW Transmembrane.
FT SIGNAL 1 33
FT PROPEP 34 187

```

FT CHAIN 188 1902 PII-TYPE PROTEINASE.
FT DOMAIN 188 1876 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1877 1895 MEMBRANE ANCHOR (POTENTIAL).
FT DOMAIN 1896 1902 CYTOPLASMIC (POTENTIAL).
FT ACT_SITE 217 217 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 281 281 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 620 620 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DOMAIN 1867 1872 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT PROTEINS.
SO SEQUENCE 1902 AA; 200550 MW; 87CECBA9345FE9D3 CRC64;

Query Match 12.6%; Score 434.5; DB 1; Length 1902;
Best Local Similarity 27.0%; Pred. No. 5,96-14;
Matches 164; Conservative 92; Mismatches 240; Indels 111; Gaps 25;

Y 135 -----YDRFNGSAIETTPDNLDKIKQIEGSSVERAQQVOPMMHARKEIGVEAID 187
Db 84 NKVQODIYVIVQMSAA-PASENGILRTDYSTFAEQETNKYIAQAQAVKAVEQVT 142
Y 78 NSNKSQGYTSPFNKNTENKKEKVV---YIAEFKKEGGERAKIKLSLKTKVLYT 134
Db 84 NKVQODIYVIVQMSAA-PASENGILRTDYSTFAEQETNKYIAQAQAVKAVEQVT 142
Y 143 QOTAGESGYVNGESTVRVVDIPKQLQAGVKTLYLAKYIYPPDAKANSNAVQAVWS 202
Db 188 YLKSINAEFGKNGFCRWVSNIDTGDYRHKAMRIDDAKASMRKKEDL-KGTD--K 243
Y 203 NYK-----YKGEYVSVISGIDPTHKMRLSD--KDYKLYTKSDVEKFTDVKH 251
Db 244 NYWLSDKIPHAFNY-NGKLTVEKYDGRDYFDPHGHINGILAGNTEDDIKNFND 302
Db 252 GRYEFSKYPGFENYADNDITTDKVD-----EQHGMHAGIIGANTGDD--PAKSV 303
Y 303 GIAPNAQIFYKMYSDAG-SGFAGDETFMFAIEDSIKHNVVSVSSGF-TGTGLVGEKY 360
Db 304 GVAPEOQLAKKAVSNSTSAKTSATVSAIEDSAKIGADVLNLSGNSNGNLTLEPE 363
Y 361 WOAIRALKKAGIPVNAVATGATVATSASSSSWDLVANNHLMKMTDGTGNTTAHEDAIAVAS 420
Db 364 LAAVONANESGTAIVISAGNSGTGSAT--EGVKNIDYGLDNEKVSFGTSRGATVAS 421
Y 421 AKNOTVERDKNI-----GGSFKY--RNIGAFPKSKITINEDCTKAPSKLFPYI 470
Db 422 AENDVITQAVTITDGTGLQLGPEITQLSSHDFTGSPQKFFYLVKQASG-----NL 473
Y 471 GKGDODLIGLDLKGKIAVMDR-IYTKDLKNAFKAMDKAGARIVVTVVYVNRDMTE 529
Db 474 SKGALADYTA-DAGKTIIVRGESFPDKQKYNQA--AGAAGLIIVT-----DSTA 523
Y 530 LPAMGYEADGTSQVSEISGDDGVKL--WNMINDPKKTVEYKRNKEDKDKLEQYYPID 587
Db 524 TPMTSIALT--TTEPTFGLSSVTGQKLVDTVAHPDDSLGVK-----563
Y 588 MESTNSNKNPNVGDEKIDFKAPDPTDKELYKEDIIVPAGSVMGPRIDLLKPOVSAPGK 647
Db 564 -----ITLAMPN---OKYTEDKM--SDFTSYGVSNLSFKRPITTAGG 602
Y 648 NIKSTLN 654
Db 603 NIMSTON 609

RESULT 2
P2P_LACPA STANDARD; PRT; 1902 AA.
ID P2P_LACPA 002470;
AC 002470;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE PII-TYPE PROTEINASE PRECURSOR (EC 3.4.21.96) (LACTOCEPIN) (CELL WALL-
DE ASSOCIATED SERINE PROTEINASE) (LP151).
DE PRP.
OS Lactobacillus paracasei.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;

```

```

OC Lactobacillus.
OX NCBI_Taxid-1597;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NCDO 151;
RX MEDLINE-92381481; PubMed-1512565;
RA Hoick A., Naes H.;
RT "Cloning, sequencing and expression of the gene encoding the cell-
RT envelope-associated proteinase from Lactobacillus paracasei subsp.
RT paracasei NCDO 151.";
RL J. Gen. Microbiol. 138:1353-1364(1992).
RN [2]
RP SEQUENCE OF 189-196.
RX MEDLINE-9222694; PubMed-1564442;
RA Naes H., Nissen-Meyer J.;
RT "Purification and N-terminal amino acid sequence determination of the
RT cell wall-bound proteinase from Lactobacillus paracasei subsp.
RT paracasei.";
RL J. Gen. Microbiol. 138:313-318(1992).
CC -1- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE
CC GROWTH OF THE BACTERIA ON MILK.
CC -1- CATALYTIC ACTIVITY: ENDOPEPTIDASE ACTIVITY WITH VERY BROAD
CC SPECIFICITY, ALTHOUGH SOME SUBSTRATE PREFERENCE HAVE BEEN NOTED.
CC E.G. LARGE HYDROPHOBIC RESIDUES IN THE P1 AND P4 POSITIONS, AND
CC PRO IN THE P2 POSITION, BEST KNOWN FOR ITS ACTION ON CASEINS, AND
CC ALTHOUGH IT HAS BEEN SHOWN TO HYDROLYSE HEMOGLOBIN AND OXIDIZED
CC INSULIN B-CHAIN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, CELL WALL.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M83946; AAA25248.1; -.
DR PIR: B4858; B4858.
DR HSP: Q99405; IMPT.
DR MEROPS: S08.019; -.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003137; PA.
DR InterPro: IPR002029; Peptidase_S8.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF02225; PA; 1.
DR Pfam: PF00082; Peptidase_S8; 3.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE HIS; 1.
DR PROSITE: PS00138; SUBTILASE SER; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; FALSE_NEG.
DR Hydrolase: Serine protease; Cell wall; zymogen; signal;
KW Transmembrane.
FT SIGNAL 1 33
FT PROPEP 34 187
FT CHAIN 188 1902
FT DOMAIN 188 1876
FT TRANSMEM 1877 1895
FT DOMAIN 1896 1902
FT ACT_SITE 217 217
FT ACT_SITE 281 281
FT ACT_SITE 620 620
FT DOMAIN 1867 1872
SO SEQUENCE 1902 AA; 200253 MW; D8C9F38CE5DA582 CRC64;

Query Match 12.4%; Score 429.5; DB 1; Length 1902;
Best Local Similarity 26.1%; Pred. No. 1e-13;
Matches 174; Conservative 97; Mismatches 275; Indels 121; Gaps 27;

```

```

OY 17 ISDKKVAENKEKHEHNIHSAMETSDQEKKTAVIKEEVNSKNPYIDNNTSNEAKIKE 76
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 35 ISQQTAVS-----SLANTVKAATAKQAAT--DTTAATTNQAALQALAKGIDYKN 82
OY 77 ENSNKSQGDYDSFVNKNTENPKED---KVYIAEFKDESEKAIKELSLKNTKVLV 133
    83 LNKVQOQDIYDVIVQMSA--PASENGTLRTDYSSTAIEQETNKYIAQASVKAVEQV 141
OY 134 T-----VDRIENGSAIETTPNLDKIKQIEGISSVERAKQVPMNHAKKEIGVEEAT 186
    142 TQQTAGESTGYVNGFSTKRVVDIPKLOIAGVKTVTLAKVYPTDAKASMANVQAVW 201
OY 187 DYLSINAPFGKNFDRGVNISNIDGTDYRHKAMRIDDAKASM--REKKEKDGTDKN- 244
    202 SNVK-----YKGGGVTVSVIDGIDPTHKDMRLSDOKDKVLKRYDEKFTDAKHG 252
OY 245 YMLSDKIPIHAFNYY--NGKRTTVEKYDDGRDYFPPHGHMIGIILAGNDEDDINFGNIDG 303
    253 RYFTSVVPYGFENYADNNDIT--DDTVD--EQHGMHVGIIIGANGTGDD--PKRSVVG 304
    304 IAPNAOIFSKMYSDAG--SGFAGDETFMFAIEDSIKHNVPVSVSGF--TGGLVGEKYM 361
    305 VAPDAOLAKKVTNNSDTSAITGSATLVSATIEDSAKIGADVLNNSLGSOSGNOTLEDPEI 364
OY 362 QAIRALRKAGIPNVVATGNVATSSASSMDLVANHLKMTDGTGNTVTAHEDAIANASA 421
    365 AAVQNNESGTAIVISAGNSGTSATQG--VAKDYGLQDNEMVGPQSRGATTVASA 422
OY 422 KNOTVEEDKYN-----GGESEFY--RNICAFEDKSKITTNEDGTAKPSKLEFYIG 471
    423 ENPDVISOAVITIDGKDLQIGPETIQLSSNDPFGSEFOKKFYVVKDASGLSK-----G 476
OY 472 KGGDDQLIGDLGKIAVMDRITTKDLKNAFK--AMDKARAIMVNTNYNNROWMTE 529
    477 AAADYT--ADAGKTAIVR---GELNPFADKRYAQAAGAGLITV-----NGTA 523
Db 530 LPAMGEADEGKTSQVFSISGDDGVKL--WNMINPDKTEVKRNKNEDEKDLQOYYPID 587
    524 TPLTSLRL--TFPTFGLSKSGOKKLVMDVTAHPDPSLGK-----563
OY 588 MESFNSKNKPVGDEKEIDFAPDTEKELYKEDIIVPAGTSNGPRIDLLKPDVSAK 647
    564 -----IALTLPLN--QKYTEDKM--SDFTSYGPAVSNLSEKPDITAPG 602
OY 648 NIKSTLN 654
    603 NIMSTON 609

```

```

CC -1- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE
CC GROWTH OF THE BACTERIA ON MILK.
CC -1- CATALYTIC ACTIVITY: ENDOPEPTIDASE ACTIVITY WITH VERY BROAD
CC SPECIFICITY, ALTHOUGH SOME SUBSITE PREFERENCE HAVE BEEN NOTED,
CC E.G. LARGE HYDROPHOBIC RESIDUES IN THE P1 AND P4 POSITIONS, AND
CC PRO IN THE P2 POSITION. BEST KNOWN FOR ITS ACTION ON CASEINS,
CC ALTHOUGH IT HAS BEEN SHOWN TO HYDROLYSE HEMOGLOBIN AND OXIDIZED
CC INSULIN B-CHAIN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. CELL WALL.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: X14130; CAA32350.1; -.
CC PIR: S06997; S06997.
CC HSSP: P00782; 2SBR.
CC MEROPS: S08.019; -.
CC InterPro: IPR001899; Gram_pos_anchor.
CC InterPro: IPR003137; PA.
CC InterPro: IPR000209; Peptidase_S8.
CC Pfam: PF00746; Gram_pos_anchor; 1.
CC Pfam: PF02225; PA; 1.
CC Pfam: PF00082; Peptidase_S8; 3.
CC DR PRINTS: PR00723; SUBTILISIN.
CC PROSITE: PS00136; SUBTILASE_ASP; 1.
CC PROSITE: PS00137; SUBTILASE_HIS; 1.
CC PROSITE: PS00138; SUBTILASE_SER; 1.
CC PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
CC K W Hydrolyase; Serine protease; Cell wall; Zymogen; Signal; Plasmid;
CC K W Transmembrane.
CC FT SIGNAL 1 33
CC FT PROPEP 34 187
CC FT CHAIN 188 1902
CC FT DOMAIN 188 1895
CC FT TRANSMEM 1877 1895
CC FT DOMAIN 1896 1902
CC FT ACT_SITE 217 217
CC FT ACT_SITE 281 281
CC FT ACT_SITE 620 620
CC FT DOMAIN 1867 1872
CC FT SEQUENCE 1902 AA; 200139 MW; 4B8DB844D88CDF7 CRC64;
CC FT PROTEINS.

```

Query Match 12.4%; Score 427.5; DB 1; Length 1902;
Best Local Similarity 27.5%; Pred. No. 1,3e-13;
Matches 167; Conservative 89; Mismatches 239; Indels 113; Gaps 26;

```

OY 78 NSNKSQGDYDSFVNKNTENPKED---KVYIAEFKDESEKAIKELSLKNTKVLV 134
    84 NKVQOQDIYDVIVQMSA--PASENGTLRTDYSSTAIEQETNKYIAQASVKAVEQV 142
OY 135 T-----VDRIENGSAIETTPNLDKIKQIEGISSVERAKQVPMNHAKKEIGVEEAT 187
    143 QQTAGESTGYVNGFSTKRVVDIPKLOIAGVKTVTLAKVYPTDAKASMANVQAVWS 202
OY 188 DYLSINAPFGKNFDRGVNISNIDGTDYRHKAMRIDDAKASMRPKED--KGT--K 243
    203 SNVK-----YKGGGVTVSVIDGIDPTHKDMRLSDOKDKVLKRYDEKFTDAKH 251
OY 244 YMLSDKIPIHAFNYY--NGKRTTVEKYDDGRDYFPPHGHMIGIILAGNDEDDINFGNID 302
    252 GRFNSKVPYGFENYADNNDIT--DDTVD--EQHGMHVGIIIGANGTGDD--PAKSV 303
OY 303 GIAPNAOIFSKMYSDAG--SGFAGDETFMFAIEDSIKHNVPVSVSGF--TGGLVGEK 360
    361 VAPDAOLAKKVTNNSDTSAITGSATLVSATIEDSAKIGADVLNNSLGSOSGNOTLEDPEI

```

```

Db 304 GVAPBAOLLMKVFPTNSDTSATGSAVLVAIEDSAKIGADVLLNLSGDSGNOITLEDPE 363
OY 361 WQAIRALKRAGIPVAVTGNATVATSSSSMDLVANNHLMKMTDGVNTAAHEDAIAVAS 420
Db 364 LAAVONANESGTAIVSAGNSTGSAAT--EGVKNQDYGLQDNEMVGTPTGSRGATTVAAS 421
OY 421 AKNOTVEFDKYN-----GGESFKY--RNIGAFDPKSKITTNEDGTAKAPSKLFEVYI 470
Db 422 AENTDVITQAVTITDGTGLQIGPFTIQLSSNDFGSPQKKFYVVKDASGNLSK----- 475
OY 471 KKGQODDILGLDKRKIAVMRIYTKDLKNAFK--AMDKARALMVNTVYVYRDWMT 528
Db 476 GKVAAYT--ADAKGKIAIVK--GELTFADKOKRYAAGAGLIIYV--NDGTAT 524
OY 529 ELPMAGYEADGTSQVFSIGDDGVKL--NMNINPDKKTEYKRNKEDFKDLQYYPPI 586
Db 525 PVTSMALT--TTPPTGLSSVYGQKLVDAVAAHPDLSLGVK----- 563
OY 567 DMESFNSKPNVGDKEIDFKFAPDTDELYKEDIIVAGSTSMGPRIDLKLPVSAAG 646
Db 564 -----IATLIVPN--QKYTEDKM--SDFTSYGPAVNSLSPKPDITAPG 601
OY 647 NIKSTLN 654
Db 602 GNIMSTON 609

RESULT 4
PIE_LACLC STANDARD: PRT: 1902 AA.
ID PIE_LACLC
AC 016271:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PI-TYPE PROTEINASE PRECURSOR (EC 3.4.21.-) (WALL-ASSOCIATED SERINE
DE PROTEINASE).
GN PRTP.
OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OG Plasmid pMV05.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
OX NCBI_Taxid=1359;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MG2;
RX MEDLINE=88149035; PubMed=3276687;
RA Kok J., Leenhouts K.J., Haandrikman A.J., Ledebor A.M., Venema G.;
RT "Nucleotide sequence of the cell wall proteinase gene of
RT Streptococcus cremoris Wg2.";
RL Appl. Environ. Microbiol. 54:231-238(1988).
CC -1- FUNCTION: PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE
CC GROWTH OF THE BACTERIA ON MILK.
CC -1- CATALYTIC ACTIVITY: ENDOPEPTIDASE ACTIVITY WITH VERY BROAD
CC SPECIFICITY, ALTHOUGH SOME SUBSITE PREFERENCE HAVE BEEN NOTED.
CC E.G. LARGE HYDROPHOBIC RESIDUES IN THE P1 AND P4 POSITIONS, AND
CC P6 IN THE P2 POSITION, BEST KNOWN FOR ITS ACTION ON CASEINS, AND
CC ALTHOUGH IT HAS BEEN SHOWN TO HYDROLYSE HEMOGLOBIN AND OXIDIZED
CC INSULIN B-CHAIN.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, CELL WALL.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M24767; AAA17677.1; -
CC HSSP: 099405; IMPT.
CC MEROPS: S08.019; -

```

```

DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003137; PA.
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF02225; PA; 1.
DR Pfam: PF00082; Peptidase_S8; 3.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; 1.
DR HydroLase: Serine protease; Cell wall; Zymogen; Signal; Plasmid;
KW Transmembrane.
FT SIGNAL 1 33
FT PROPEP 34 187
FT CHAIN 188 1902
FT DOMAIN 188 1876
FT TRANSMEM 1877 1895
FT DOMAIN 1896 1902
FT ACT_SITE 217 217
FT ACT_SITE 281 281
FT ACT_SITE 620 620
FT DOMAIN 1867 1872
FT SEQUENCE 1902 AA; 199910 MW; 2901C7F19B2E5D0B CRC64;

Query Match 12.28; Score 421.5; DB 1; Length 1902;
Best Local Similarity 26.98; Pred. No. 2.5e-13;
Matches 163; Conservative 89; Mismatches 244; Indels 111; Gaps 25;

OY 78 NSNKSQGDYTSFVNKNTENPKED--KVVYIAEFKDGEGEKAIRKELSLKTKTKVLT 134
Db 84 NKVQODIYVAVIOMSA--PASENGTLRTYSSAEIQGTNNVIAQAQVKAAYEVY 142
OY 135 -----YRIENGSALETPDNLKIKQIEGSSVERAQVQPMNHAREIGVEAID 187
Db 143 QQTAGESEYGVVNGFSTKRVAVDIPKLIQIAGVTVLAKYVPTDAKANSMAVQAQVMS 202
OY 188 YLKSINAPFGKNPGRGVVSNIDGTCDYRKARIDDDAKASRFKEDL-KETD--K 243
Db 203 NYK-----YKGEGVSVSIDGIDPTHKDKRLSD--KDVKLRKSDVEKFTDAKH 251
OY 244 NYWLSDKIPHAFNYY-NGKITYVEKDDGRDYPHGHMIGIAGNDEIDIKNFNCID 302
Db 252 GRVNSKVPYGFNADNDIT--DDTVV--EQHGHNVAIGIANGTGD--PAKSV 303
OY 303 GIAPNAQIFSYKMYSDAG-SGFAGDETFHAIEDSIKHNADVSVSSGF-TGTGLVGEKY 360
Db 304 GVAPBAOLLMKVFPTNSDTSATGSAVLVAIEDSAKIGADVLLNLSGDSGNOITLEDPE 363
OY 361 WQAIRALKRAGIPVAVTGNATVATSSSSMDLVANNHLMKMTDGVNTAAHEDAIAVAS 420
Db 364 LAAVONANESGTAIVSAGNSTGSAAT--EGVKNQDYGLQDNEMVGTPTGSRGATTVAAS 421
OY 421 AKNOTVEFDKYN-----GGESFKYRNIGAFDPKSKITTNEDGTAKAPSKLFEVYI 470
Db 422 AENTDVITQAVTITDGTGLQIGPFTIQLSSNDFGSPQKKFYVVKDASGNLSK----- 475
OY 471 KKGQODDILGLDKRKIAVMRIYTKDLKNAFK--AMDKARALMVNTVYVYRDWMT 528
Db 474 SKGALADYTA-DAKAKIAIVRGELSPDDKOKYAQA--AGAAGLIIV--NDGTATP 525
OY 530 LPAMGYEADGTSQVFSIGDDGVKL--NMNINPDKKTEYKRNKEDFKDLQYYPID 587
Db 526 VTSMAIT--TTPPTGLSSVYGQKLVDAVAAHPDLSLGVK----- 563
OY 588 MESFNSKPNVGDKEIDFKFAPDTDELYKEDIIVAGSTSMGPRIDLKLPVSAAG 647
Db 564 -----IATLIVPN--QKYTEDKM--SDFTSYGPAVNSLSPKPDITAPG 602
OY 648 NIKSTLN 654

```



```

CC -1- FUNCTION: THIS VIRULENCE FACTOR OF S.PYOGENES SPECIFICALLY CLEAVES
CC THE HUMAN SERUM CHEMOTAXIN C5A AT LYS(68)-ASP(69) BOND NEAR ITS
CC C-TERMINUS, DESTROYING ITS ABILITY TO SERVE AS A CHEMOATTRACTANT.
CC -1- SUBCELLULAR LOCATION: CELL WALL BOUND.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J05229; AAA26960.1; -.
DR PIR: A35066; A35066.
DR HSSP: P00782; 2S8T.
DR MEROPS: S08.020; -.
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003137; PA.
DR InterPro: IPR002029; Peptidase_S8.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF02225; PA; 1.
DR Pfam: PF00082; Peptidase_S8; 4.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; FALSE_NEG.
DR HydroLase: Serine protease; Signal; Cell wall; Repeat; Transmembrane.
KW SIGNAL
FT CHAIN 1 31
FT DOMAIN 32 1167 C5A PEPTIDASE.
FT TRANSSEM 32 1139 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 1140 1157 POTENTIAL.
FT DOMAIN 1158 1167 CYTOPLASMIC (POTENTIAL).
FT ACT_SITE 130 130 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 193 193 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 512 512 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT DOMAIN 1029 1104 CELL WALL ATTACHMENT DOMAIN (POTENTIAL).
FT DOMAIN 1034 1101 4 X 17 AA TANDEM REPEATS.
FT REPEAT 1034 1050 1.
FT REPEAT 1051 1067 2.
FT REPEAT 1068 1084 3.
FT REPEAT 1085 1101 4.
FT DOMAIN 1127 1131 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
FT PROTEINS.
SQ SEQUENCE 1167 AA; 128263 MW; D2DDC52E5752DA5D CRC64;

Query Match 8.5%; Score 293; DB 1; Length 1167;
Best Local Similarity 23.0%; Pred. No. 2,7e-07;
Matches 139; Conservative 99; Mismatches 207; Indels 166; Gaps 29;

```

```

OY 373 PMVATGCVATSSASSSSMDLVANNHLMKMTDGTGNTYRTAAHEDATAVA--SAKNOTVERDK 430
OY 288 SIVSAGNDSSFGGKTRPLA--DH--PDGCVGTTPAAASTILVASISPDKDLTERAM 342
OY 431 VNIGSESEKRYNNIAFPDKSKITTNEGDTKAPSK-LKFVYIGKODDILDLRGKIAV 489
OY 343 VKTDQDQDK-----EMPLSTNR--PEPNKAYDYAVANRQMKEDDF-KDVGKIAL 390
OY 490 MDRIYTKOLNAFKKAMDKGARALMVNTVNYVRD-----NMTELPMAGYEADSGT 541
OY 391 IER-GDIDFKKVKANAKKAGAVLI-----YDNDGKFPLEPVDQMPAA----- 436
OY 542 KSQVFSISGDDGKLVMMNINPDKTEVRKNNKEDFKLEQYYPIDMEFSNKNPVNGE 601
OY 437 -----FISRKDGILLKD--NP-----Q 451
OY 602 KEIDFKFAPDPTDKELYEDITVPAGS-----TSMGPRIIDLKPDVSAFGKNIKSTLN 654
OY 452 KTIFFNATPK-----VLPTASGTRKLSRFSWGLTADGNIKPIDAAPGODILS--S 499
OY 655 VINGK 659
OY 500 VANNK 504

RESULT 7
SCAP2_STREPY STANDARD: PRT: 1161 AA.
ID PS6099;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE C5A PEPTIDASE PRECURSOR (EC 3.4.21.-) (Scp).
GN SCPA OR SPY2010.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; PubMed=1126296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yvan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus
RT pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -1- FUNCTION: THIS VIRULENCE FACTOR OF S.PYOGENES SPECIFICALLY CLEAVES
CC THE HUMAN SERUM CHEMOTAXIN C5A AT LYS(68)-ASP(69) BOND NEAR ITS
CC C-TERMINUS, DESTROYING ITS ABILITY TO SERVE AS A CHEMOATTRACTANT.
CC -1- SUBCELLULAR LOCATION: CELL WALL BOUND.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE006623; AK34691.1; -.
DR MEROPS: S08.020; -.
DR PROSITE: PS00136; SUBTILASE ASP; 1.
DR PROSITE: PS00137; SUBTILASE HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW HydroLase: Serine protease; Signal; Cell wall; Repeat; Transmembrane;
KW Complete proteome.
FT SIGNAL 1 31 BY SIMILARITY.

```

```

FT CHAIN 32 1181 C5A PEPTIDASE.
FT DOMAIN 32 1156 EXTRACELLULAR (POTENTIAL).
FT TRAMSEM 1157 1174 POTENTIAL.
FT DOMAIN 1175 1181 CYTOPLASMIC (POTENTIAL).
FT ACT_SITE 130 130 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 133 193 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 512 512 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DOMAIN 1029 1121 CELL WALL ATTACHMENT DOMAIN (POTENTIAL).
FT DOMAIN 1034 1118 5 X 17 AA TANDEN REPEATS.
FT REPEAT 1034 1050 1.
FT REPEAT 1051 1067 2.
FT REPEAT 1068 1084 3.
FT REPEAT 1085 1101 4.
FT REPEAT 1102 1118 5.
FT DOMAIN 1144 1148 CONSERVED IN GRAM-POSITIVE COCCI SURFACE
PROTEINS.
SQ SEQUENCE 1181 AA; 129486 MM; 3F9FC51763419CFC CRC64;

```

Query Match 8.5%; Score 292; DB 1; Length 1181;

Best Local Similarity 23.0%; Pred. No. 3.1e-07;

Matches 139; Conservative 93; Mismatches 207; Indels 166; Gaps 29;

```

QY 97 NPKKEDKVVYIAEFKDESGEKAI-----KELSLKNTKLVLYYDRINGSAIETT 147
DB 24 MMSDIAKANTVTE--DTPATEQAVETPOPTAVSEAPSSKETKTPTPD-----DAEETI 76
QY 148 PDLKDKIKIEGSSVERAKQVPMNHAKKEIGVEALD--YLKSLNAPFGKPFGRGM 205
DB 77 ADANLAPQAPAKTAD-----TPATSKATYIRDLPDSOVKTLQEKAGK---GAGT 124
QY 206 VISNIDTGTDRHKAMRIDDAKASMRFKEDLGTGDKNY-----WLSDKIPHAENYNN 259
DB 125 VAVVIDAGDKNNEARLDKTKARYQ--SKEDLEKAKKEGITYGSEVNDKVAIYHDSK 183
QY 260 GKKTYEKKDGDGNDYDPPHGMHITAGLAGNDTEQDITKNFNIGDIAIPNAIESYKK---- 315
DB 184 DGTAVDQ-----EHGTHVSGILSGN--APSETKEPYRLEGAMPQQLLMRVEIVN 233
QY 316 -VSDAGSGAGDETMFHAIEDSIKHNVDVSVSSGFTGGLVG--EKYQAIATLAKKAGI 372
DB 234 GLADYARNTA-----QALIDAVNLCAKAYINMFGNAAALAYANLPBETKKAEPYASKGV 287
QY 373 PMVATGNATYATSSSSWDLVANNHLKMTDGNVTRTAHEDAIAVA--SAKQVTEFDK 430
DB 288 SITVSAGNDSSEGGKTRPLA--DH---PDYGVVGTTPAADSTLYTAVASPDQILLETAT 342
QY 431 VNIGSEFYRNIGAFDPKSKITTNEDGTAKPSK-LKPYVIGGQDODLIGDLRGKIAV 489
DB 343 VKTADQODK-----EMPVILSTNR---FEPYKAYDVAVANRGMKEDDF--KDYKGIATL 390
QY 490 MDRIYRKDLKNAFKKAMDGARIMVNVNYYNRD-----NMTELPRAMQYEADEGT 541
DB 391 IER-GIDTRDKTAAKKAGAVGLT---YDNQDKFPIELPNVDQMAA----- 436
QY 542 KSGVFISGDDGVKLMNMINPKKTEVKRNKKEPDFDKLEQVYPIDESEFNSNKPVNGDE 601
DB 437 -----FISKKDGILL-----KEN-----PQ 451
QY 602 KEIDFPAEDTKELYKEDIYVAGS-----TSWGPRIIDLKLPDVSAPGNIKSTLN 654
DB 452 KTTTFNAATRK-----VLPATAGTKLSRFSWGLTADGNIKPDIAAPGODILS--S 499
QY 655 VINGK 659
DB 500 VANNK 504

```

RESULT 8
SUBF_BACSU STANDARD; PRT; 1433 AA.
AC P16397;
DT 01-AUG-1990 (rel. 15, Created)

```

DT 01-FEB-1996 (rel. 33, last sequence update)
DT 20-AUG-2001 (rel. 40, last annotation update)
DE BACTILOPEPTIDASE F PRECURSOR (EC 3.4.21.-) (ESTERASE) (RP-I PROTEASE)
DE (90 KDA SERINE PROTEINASE).
GN BPR OR BPF.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 195-222.
RX MEDLINE=90170864; PubMed=2106512;
RA Sloma A., Rufo G.A. Jr., Rudolph C.F., Sullivan B.J., Theriault K.A.,
RA Pero J.;
RT "Bacillopeptidase F of Bacillus subtilis: purification of the protein
RT and cloning of the gene.";
RL J. Bacteriol. 172:1470-1477(1990).
RN [2]
RP REVISIONS.
RC STRAIN=168;
RX MEDLINE=90368623; PubMed=2118514;
RA Sloma A., Rufo G.A. Jr., Rudolph C.F., Sullivan B.J., Theriault K.A.,
RA Pero J.;
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=90216713; PubMed=2108961;
RA Wu X.-C., Nathoo S., Pang A.S.-H., Carne T., Wang S.-L.;
RT "Cloning, genetic organization, and characterization of a structural
RT gene encoding bacillopeptidase F from Bacillus subtilis.";
RL J. Biol. Chem. 265:6845-6850(1990).
RN [4]
RP SEQUENCE OF 1-211 FROM N.A.
RX MEDLINE=89008108; PubMed=3139638;
RA Beall B., Lowe M., Lutkenhaus J.;
RT "Cloning and characterization of Bacillus subtilis homologs of
RT Escherichia coli cell division genesftsZ andftsA.";
RL J. Bacteriol. 170:4855-4864(1988).
RN [5]
RP SEQUENCE OF 1410-1433 FROM N.A.
RX STRAIN=168 / MARBURG;
RX MEDLINE=90174995; PubMed=2106671;
RA Masuda E.S., Anaguchi H., Sato T., Takeuchi M., Kobayashi Y.;
RT "Nucleotide sequence of the sporulation gene spoIIIGA from Bacillus
RT subtilis.";
RL Nucleic Acids Res. 18:657-657(1990).
RN [6]
RP SEQUENCE OF 195-219.
RC STRAIN=NATTO 16;
RA Kato T., Yamagata Y., Arai T., Ichishima E.;
RT "Purification of a new extracellular 90-kDa serine proteinase with
RT isoelectric point of 3.9 from Bacillus subtilis (natto) and
RT elucidation of its distinct mode of action.";
RL Biosci. Biotechnol. Biochem. 56:1166-1168(1992).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
SUBTILASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sb.ch/announce/
CC or send an email to license@isb-sb.ch).
CC -----
CC EMBL; M29035; AAA62679.1; -
CC EMBL; J05400; AAA83362.1; -
CC EMBL; M22630; AAA22458.1; -
CC EMBL; X17344; CAA35224.1; -
CC EMBL; Z89111; CAB13403.1; -
CC EMBL; Z89112; CAB13404.1; -
CC PIR; A35131; A35131.

```


DR PIR: A35750; A35750.
 DR PIR: A36734; A36734.
 DR HSP: P00782; 2SPT.
 DR MEROPS: S08_017; -.
 DR Subtilisin; Bg10233; bpt.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; FALSE_NEG.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 DR Hydrolase; Serine protease; 1.
 KW Hydrolase; Serine protease; 1.
 FT SIGNAL 1 30
 FT PROPEP 31 194
 FT CHAIN 195 755
 FT PROPEP 756 1433
 FT ACT_SITE 227 227
 FT ACT_SITE 274 274
 FT ACT_SITE 452 452
 FT CONFLICT 219 219
 FT CONFLICT 393 393
 FT CONFLICT 829 834
 FT CONFLICT 836 841
 FT CONFLICT 844 852
 FT CONFLICT 853 1433
 SQ SEQUENCE 1433 AA; 154577 MW; 98DF6846897807C9 CRC64;

Query Match 6.98; Score 239.5; DB 1; Length 1433;
 Best Local Similarity 20.58; Pred. No. 0.00015;
 Matches 156; Conservative 128; Mismatches 257; Indels 221; Gaps 37;

10 SDEALNIDSKAEVNEKEHENHSAETQDFEKKTVIKKEVYKRPVINDNTSN 69
 30 ASSKVTGSYVKEKQSAESIONKISSSLKSKFKKEKTFLEIKRKDLA--NPEAKAAAY 87
 70 EBAKIEBNSKSGQDYTFVFN--KNTENPKKEDKVYVIAEFKDESEKAIEKLSLK 127
 88 KKAKSKKLSAKTEQKRSAYVSSLKVTADSEQDYKLTQDKGNADQ----- 138
 128 NTKVLYYDRFNGSAIETTDPNDIKIQIGISSEVAQVQ-----PM-MHARKEI 180
 139 ---IHST-YVANGIAVHASKVEKVEKVVQFEVEKVLNPEKROLEKSSSPNMKKAOKAI 193
 181 GVEAIDY-LKSIAP--FGKFDGRGVINIDPTGTIRKAKRIDDAKMRKKED 237
 194 KATGGEVNNVQIDAPKAMALGDTGTGVVASIDTGVAMNIPAL-----KEK 240
 238 LKGTDKNWLSDKIPHAENYNGKRTVEKYDGRDYFDPGHMIIAGIACNDTQDIKN 297
 241 YRGYNPEN--PNEPENENMNYDAVAGEASPYD-----LAHGTHTGTMGVSEPD---G 289
 298 FNGIDGIAPNQIFSYKMYSDAGSG---FAGDETMAHLEDISKH---NVDVVSVSSG 349
 290 TNOI-GVAPGAKMTAVAKFSEDSGTDADILEAGEVNLAPKDAEENPHEMAPDVVNNWG 348
 350 FTGTGLGEEKYQWQIRALRKAGIPVAVATGN-----VATSSA-- 386
 349 -GGSGL-DEWYRDVNMNMRADIFEPFSAGNTDLFIPGGPGSIANPANYPESFATGADI 406
 387 -----SSWD--LVANNHLK---MTD 402
 407 NKKLADESLQSPSYDEIKPEISAPGVNIRSSVPQGYEDMDGTSMGPHVSAVALLK 466
 403 TGNVTRIAHEDATAVASAKNQYVEF--DKVNIAGESKRYKNIGAFPKSKITINEDCT- 459
 467 QANASLSVDEMEDILTSTAEPSTDPDSPNG---YGHVLNAPFAVASAVT---DELG 520
 460 KAPKELKVVYIGKQDOD-----LIGDLRGKIIVMRIRYTKDKNAFKKMD 507
 521 KAEQVSV-----EGDDQEPYVQHEKVTAEVGEGLPTLTAEDNVSTYSKLSYK--LD 574
 508 KGARAIWVNTVNYNRDNMTLPAKMGVEADGTRKSOVFSISGDDGVKLMNMINPDKTE 567

DB 575 QG-----EMTEIRAK-----RISCHLAKTYQAEIPIDK-- 603
 QY 568 VKRNKKEDEFRKLEQY--YPIDMESFN-SNKPNGDEKEIDFKRAPDIDRELYEDIYVP 624
 DB 604 --GTRKSLKWMHIDFGGHVSSDYDYTVKPSITAGYKQDFETAPG-----WVA 651
 QY 625 AGSTS--WGPRIIDLKPDVASAPCKNIKISPLNVIKSGSTYG 663
 DB 652 SCTNNNMWNG-----VPSYGPN-----TASGKXYVG 678
 RESULT 9
 WPIA_BACSU STANDARD; PRT: 894 AA.
 AC P54423; 006726;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE CELL WALL-ASSOCIATED PROTEASE PRECURSOR (EC 3.4.21.-) [CONTAINS: CELL WALL-ASSOCIATED POLYPEPTIDES CWPB23 AND CWPB52].
 GN WPIA.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/staphylococcus group; Bacillus.
 OX NCBI_Taxid=1423;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 32-54 AND 414-428.
 RC STRAIN=168;
 RX MEDLINE=97158234; PubMed=9004506;
 RA Margot P., Karamata D.;
 RT "The wprA gene of Bacillus subtilis 168, expressed during exponential growth, encodes a cell-wall-associated protease.";
 RT microbiology 142:3437-3444(1996).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=98015415; PubMed=9353931;
 RA Medina N., Vannier F., Roche B., Autret S., Levine A., Seror S.J.;
 RT "Sequencing of regions downstream of addA (98 degrees) and cItg (289 degrees) in Bacillus subtilis.";
 RT microbiology 143:3305-3308(1997).
 CC -1- FUNCTION: NOT YET KNOWN; COULD BE INVOLVED IN PROTEOLYTICAN
 CC DEGRADATION, BY CLEAVAGE OF ITS PEPTIDE BRIDGES.
 CC -1- SUBCELLULAR LOCATION: CELL-WALL BOUND.
 CC -1- PFM: PROCESSED INTO CWPB23 AND CWPB52.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
 CC SUBTILASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U58981; AAC25926.1; -
 DR EMBL: Y09476; CAAT0641.1; -
 DR EMBL: Z99109; CAB12917.1; -
 DR HSP: Q99405; 1MPT.
 DR MEROPS: S08_004; -.
 DR Subtilisin; Bg11846; WPIA.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; FALSE_NEG.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Hydrolase; Serine protease; Cell wall; zymogen; signal;
 KW Complete proteome.
 FT SIGNAL 1 31
 FT CHAIN 32 894
 FT CHAIN 894 894
 FT CWPB23.
 FT CELL WALL-ASSOCIATED PROTEASE.

FT PROBE 413 POTENTIAL.
 FT ACT 414 894 CMBP52.
 FT ACT SITE 466 466 RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 497 497 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 650 650 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT CONFLICT 9 9 V -> A (TN REF. 1).
 FT CONFLICT 14 14 L -> I (TN REF. 1).
 SQ SEQUENCE 894 AA: 96487 MW: 0F67C353E5FBD8C CRC64;

Query Match 5.8%; Score 201.5; DB 1; Length 894;
 Best Local Similarity 19.8%; Pred. No. 0.0059;
 Matches 155; Conservative 112; Mismatches 247; Indels 269; Gaps 40;

5 LADTSSSEDALNISDK-----EKAENKEKHNIHSAMETSO-----DFKE---KRTAVI 51
 200 LSGTAGSKELTSLYKAAPIYAKLANKARREITYODLTLPLEDDVSENGASSSYKV 259
 52 KEKEVVSKNPVIDNNTSNEBAKIEE-----NSKSGDYTDSFVNKNTENPKKE-- 101
 260 TEKQKAINRLYKALQSVSPFLKEEIKQADRLNMKQLOGKTAGAILTENNTIAKSEVQ 319
 102 -DKVYVIA-----EFKDESEKAIKELSLKNTKVLTYDRIF----- 139
 320 TTKVIRKVKDNKSLSSVHNEMKGFSAOSKDISNVKAKKL--FDNLXSELPDEKO 377
 140 NGS-----AIEETPNLDIKIQIEGISSYER---AQKVQ-----PMNHARKETGVEE 184
 378 NGVYTSASAKRVKSAAMTILSKMSNVEFAEPYQEKSLANDQYQYQWPLKNGENG--GVKN 436
 185 AIDYKLSINAPFEGKNFDRGCMVISNIDTGTDYHKKAMRIDDAKASMRKEDLKG--T 241
 437 ADVKYEPAWTLTSKR--KLNDTLIAVDTGVD-----STLADLKGKVRT 478
 242 DKVYWLSDKIPHAFNYYNGKIVYKDYDRODFDPHGMHIIAGLNGDDEQIKNNGI 301
 479 D-----LGHNFVGRNNAMP-----DQG-----HGHVAGITAAQS-----DNQYSM 515
 302 DGIAPNAQIFSYKMYSDAGSG-----FAGDE-----TMFHAIEDSIKHN 340
 516 TGINAKAKIIPVAVLDSAGSDTEQIALGIKYADKAKAVINISLGGYSRVLEFALKYA 575
 341 VD---VVSVSSGFTGTGLV---GEKYQAIIRALRK-----AGIP 373
 576 ADKNVILAAASGDGENALSYPASSKYVMSVGATNRMDTADFSNYGKGLDISAPGSDIP 635
 374 MYVATGN-----YATSSASSMDLVANNHUKMTDGNVTRTAHEDAIANASA 421
 636 SLVPNGNVYTMGSTMATPYAAAAGLLF--AQNPKLKRTVEEDMKKTADD----- 685
 422 KNOTVEFDKVNIGESFVKYNIGAFDPSKITTNEDGTAKPSKLKFEYIGKODDLIGL 481
 686 -----ISFESVD--GGEELIYDYG---DPIEI-----PKPI-----GV 714
 482 DLR---GKIAVMDRIYTKDLKNAFKKAMDGARAINVNTVNYNRDNTWELPAMGYEAD 538
 715 DMHSGYGRILVMKAVASADLQ-----LKVNKLES-----TQTAVRG--SAK 753
 539 EGRKSOVFSISGDDGVKLMNMIMPDKTEYKRNKNEDEFKLEQIYPIIDMESNSKNPNV 598
 754 EGT-----LIEVNMNGKRLKLSAKAGKD-----NAFVNI 782
 599 GDEKE---IDFKFAPDTPDELYKEDIIVPAGSYSWGPRIDLKLPVSAAPGK--NIKSTLN 654
 783 ATQKQOQOVLTKATKGDAKISTYK--VYVYVKGKSGTGPVNAVATKTQTAVAGKANSKAMIR 840
 655 VIN 657
 841 VKN 843

RESULT 10
 PRTB_YEAST

ID PRTB_YEAST STANDARD: PRT: 635 AA.
 AC P09232;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE CEREYISIN PRECURSOR (EC 3.4.21.48) (VACUOLAR PROTEASE B) (PROTEINASE
 YSCB).
 GN PRB1 OR YEL060C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycos.
 NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 281-295.
 RC STRAIN=AB320;
 RX MEDLINE=88142830; PubMed=3325823;
 RA Moehle C.M., Rizard R., Lemmon S.K., Smart J., Jones E.W.;
 RT "Protease B of the lysosomal vacuole of the yeast Saccharomyces
 RT cerevisiae is homologous to the subtilisin family of serine
 RT proteases.";
 RT Mol. Cell. Biol. 7:4390-4399(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Allen E., Araujo R.,
 RA Aviles E., Berno A., Brennan T., Carpenter J., Chen E., Cherry J.M.,
 RA Chung E., Duncan M., Guzman E., Hartzell G., Hunkeler-Smith S.,
 RA Hyman R., Kayser A., Komp C., Laskar D., Lew H., Lin D.,
 RA Mosedale D., Nakahara K., Nemach A., Norgren R., Oelner P., Oh C.,
 RA Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T., Smith V.,
 RA Taylor P., Wei Y., Yellon M., Botstein D., Davis R.W.;
 RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 605-635 FROM N.A.
 RC STRAIN=S288C;
 RA Saunders W.S., He L., Loo K.K., Hoyt M.;
 RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP PROCESSING.
 RX MEDLINE=92078140; PubMed=1744078;
 RA Nebes V.L., Jones E.W.;
 RT "Activation of the proteinase B precursor of the yeast Saccharomyces
 RT cerevisiae by autocatalysis and by an internal sequence.";
 RT J. Biol. Chem. 266:22851-22857(1991).
 CC -1- FUNCTION: AMONG OTHER SUBSTRATES, ACT ON YSCA (PEP4) TO ACTIVATE
 CC IT BY PROCESSING ITS PRO-PEPTIDE.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS WITH BROAD SPECIFICITY,
 CC AND OF BZ-ARG-OET > AC-TYR-OET. DOES NOT HYDROLYSE PEPTIDE AMIDES.
 CC -1- SUBCELLULAR LOCATION: LYSOSOME-LIKE VACUOLES.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
 CC SUBTILASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M18097; AAA34901.1; -
 CC EMBL: U18795; AAB65027.1; -
 CC EMBL: U11859; CAA77866.1; -
 CC EMBL: M80522; AAA34495.1; -
 CC PIR: A29358; A29358.
 CC HSSP: P06873; 3PRK.
 CC MEROPS: S08.052; -
 CC SGD: S0000786; PRB1.
 CC InterPro: IPR000209; Peptidase_S8.
 CC Pfam: PF00082; Peptidase_S8; 1.
 CC PRINTS: PR00723; SUBTILISIN.
 CC PROSITE: PS00136; SUBTILASE_ASP; 1.
 CC PROSITE: PS00137; SUBTILASE_HIS; 1.

DR PROSITE: PS00138: SUBTILASE SER: 1.
 KW Hydrolyase: Serine protease: Glycoprotein: zymogen: signal.
 FT SIGNAL 19
 FT PROPER 20 280
 FT CHAIN 281 635
 FT DISULFID 460 491
 FT ACT_SITE 325 325
 FT ACT_SITE 357 357
 FT ACT_SITE 519 519
 FT CARBOHYD 594 594
 FT CONFLICT 622 622
 SQ SEQUENCE 635 AA: 69621 MW: C346C2B1C7DDC48 CRC64;

Query Match 5.68; Score 194; DB 1; Length 635;
 Best Local Similarity 19.18; Pred. No. 0.0088;
 Matches 121; Conservative 90; Mismatches 207; Indels 214; Gaps 27;

QY 21 EKVAENKREHNSAMETSDQFEKKTAIVKEKVSKNPIDNNTSNEAKIKESNSN 80
 b 149 EKTLEKGNHNRRLAPLVSTAPNDPAISKITPNRIIT---YFKGAPQGEIDPIKEVQ 204
 QY 81 KSGQDYDSEVKNKNTENPKEDKVYVIAEFKDESGEKAIELSLKTKVLYTYDRIEN 140
 Db 205 QAO-----LQSEVNLASD-----AFISTKDTSLSTEAGGIDDS---FNIDNLS 248
 QY 141 GSAIETPPNDLKIKGISISVERAKVQPMNNHARKEIVEEALDYLSINAPFG--- 197
 Db 249 GYIGFTYEIDLRKQNPPLVPEVD-----SIVETPDTONSAPWGLAR 294
 QY 198 -----KNFD-----GRGVISNIDTGYDHRKAMRIDDAKA---SMRFEKE 236
 Db 295 ISHRENLGSEFNKLYDDDAGRGVTSYIDTVGNINIK---DEKKAINGKITPLNDE 350
 QY 237 DLKGTDKMYLSDKIPAFNNTNGSKITVEKYDGRDYFDPHGMHIALGNTDQDIK 296
 Db 351 DLDG-----NG-----HGTCACTIAS-----K 368
 QY 297 NFNIGIDGAPNAQIFSYKMSDAGSGFAGDETFMFAIDSIK--HNVDDVSVSSGPTGT-- 353
 Db 369 HY-----GAAKANNVAVVAVLSNNGSGTMSD--VVKGVYAAKAHOKKEKKKFGKSTA 422
 QY 354 --GLVEK--YMOAIRALRRAGIPMVAVATGNVATSSASMDLVANNHLMKMTDGNVTR 408
 Db 423 NMSLGGGKSPALDLAVNNAVEVGIHFAVAGNENDACNTS----- 463
 QY 409 TAAHEDATAVASAKNQIYEPDK-----VNIGGESFKYRNIGAFPDKSKI 452
 Db 464 PASADKALTVGAS---TLSDRAVFSNMKGVDFAPGLNI-----LSTYIGSDA 511
 QY 453 TTNEGDTKAPSK-----LKFVYIGKGODDOLIGLGRKIAVMDRIYTKDLKNAFKAM 506
 Db 512 TATLSGTSMASPHVAGLITFLSLQPGSDSEFFELG-----QDSLTPQDLK----- 558
 QY 507 DKGARAIWVTVNYNBDNNTTELPRMGYEADGTRQGVFSISGDDGVKLLNMNIPDKKT 566
 Db 559 -----KLHISTKDILEDIP-----EDTPNVLIYNGGODLSAFNN----- 594
 QY 567 EVKRNKKEDEPKLEQYPIFIMESFNKNPV 598
 Db 595 DTKRSHSGFKQEL-----NMDEFISGKIDL 620

RESULT 11
 THES_BACSP STANDARD: PRT: 401 AA.
 AC Q45670;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DE THERMOPHILIC SERINE PROTEINASE PRECURSOR (EC 3.4.21.-) (AK.1
 DE PROTEASE).
 OS Bacillus sp. (strain AK1).

OC Bacteria: Firmicutes: Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 OC NCBI_TaxID-1409;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-95085262; PubMed-7993087;
 RA Maciver B., Mchale R.H., Saul D.J., Bergquist P.L.;
 RT "Cloning and sequencing of a serine protease gene from a
 thermophilic Bacillus species and its expression in Escherichia
 coli";
 RT Appl. Environ. Microbiol. 60:3981-3988(1994).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (1.80 ANGSTROMS) OF 122-401.
 RX MEDLINE-20057863; PubMed-10588904;
 RA Smith C.A., Toogood H.S., Baker H.M., Daniel R.M., Baker E.N.;
 RT "Calcium-mediated thermostability in the subtilisin superfamily: the
 crystal structure of Bacillus Ak.1 protease at 1.8-A resolution";
 RL J. Mol. Biol. 294:1027-1040(1999).
 CC -1- COFACTOR: CALCIUM-DEPENDENT. BINDS 4 CALCIUM IONS.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- MISCELLANEOUS: HAS A PH OPTIMUM OF 8.5, A TEMPERATURE OPTIMUM OF
 75 DEGREES CELSIUS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
 SUBTILASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L29506; AAA63688.1; -;
 CC PDB: 1DBT; 18-NOV-99.
 DR DR
 DR MEROPS: S08.009; -;
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PRO0723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; 1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KW Hydrolyase: Serine protease: zymogen: signal; Calcium-binding;
 KW 3D-structure.
 FT SIGNAL 1 24
 FT PROPER 25 121
 FT CHAIN 122 401
 FT ACT_SITE 160 160
 FT ACT_SITE 193 193
 FT ACT_SITE 347 347
 FT DISULFID 258 260
 SQ SEQUENCE 401 AA: 42835 MW: 1C736EFA489F256F CRC64;

Query Match 5.58; Score 191; DB 1; Length 401;
 Best Local Similarity 21.75; Pred. No. 0.0068;
 Matches 107; Conservative 61; Mismatches 130; Indels 196; Gaps 23;

QY 51 IKEKEVSKN-----PVIDNNTSNEAK-----IKENSKOGQVTDTSFVKNKNTNP 98
 Db 1 MKFAIVSLAVSMLSPFLVEASNDGVSPKTVSEINVSHEKGAVQGEV----- 53
 QY 99 KKEDKVYIAEFKDESGEKAIEKLSLKNKTVLYTYDRIENGSALIEETPNLDKIK--- 155
 Db 54 -----IVQFKEDVNAEEKAKALKEVGATAV-----PDN-DRVKEKF 88
 QY 156 --OIEGISVERAKVQPMNNHARKEIGVEEALDYLSINAPFGKN--FDG----- 202
 Db 89 NVLKVGNVEAVKALNNPPLVEYAE-----NYL--FNAAMPNDTYGYGYQVGPON 138
 QY 203 -----RGMVSNIDTGYDHRKAMRIDDAKASMRKEDLKTGDKMYLWSD 249
 Db 139 TYTDYAMDVTKSGSGQELAVIDTGVDTYH-----PDLGK----- 173

```

QY 250 KIPAEYNGGKTTVEKDYDGRDYP-----HGMHIALGACNDTEODIKNFNGIDGI 304
D 174 -----VIKGYDFVNDYDPMDLNNGHTVAGIAAA-----ETNNATGIAGM 214
QY 305 APNAQISYKMYSDAGSGFADETMFHAIEDSIKHNVD-----VYSVSSG---FTGTGYG 357
D 215 APNTRILAVRLDRNGSGTSLD-----IADAILYADSGAEVILNLSGDCCHTTT----- 264
QY 358 EKYQAIRALAKACIPMVVATGNYATG-----ASSSSW----- 390
D 265 --LENAVYVAMNKSYYVAAAGNNGSSTTEPPASYENVYANGVADYDRLASFSGYVW 322
QY 391 DLVA-----NNHLKMTDGTGNTRTAHEDAIAVASAKN-----QTV--FDKVN 432
D 323 DVNAPGVDIYSTIGNRVAYSGTSMASPHVAGLALLASGRNNIEIRQAIETQADKIS 382
QY 433 IGSEFMYRNGAF 446
D 383 GTGYFMYGRINSY 396

RESULT 12
SUBN_BACNA STANDARD; PRT; 381 AA.
AC P35835;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SUBTILISIN NAT PRECURSOR (EC 3.4.21.62).
GN APRN.
OS Bacillus subtilis var. natto.
OC Bacillus; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86029;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-NC2-1;
RX MEDLINE-93113095; PubMed-1369081;
RA Nakamura T., Yamagata Y., Ichishima E.;
RT "Nucleotide sequence of the subtilisin NAT gene, aprN, of Bacillus
RT subtilis (natto).";
RL Biosci. Biotechnol. Biochem. 56:1869-1871(1992).
CC -1- FUNCTION: SUBTILISIN IS AN EXTRACELLULAR ALKALINE SERINE PROTEASE,
CC IT CATALYZES THE HYDROLYSIS OF PROTEINS AND PEPTIDE AMIDES.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS WITH BROAD SPECIFICITY
CC FOR PEPTIDE BONDS, AND A PREFERENCE FOR A LARGE UNCHARGED RESIDUE
CC IN P1. HYDROLYSES PEPTIDE AMIDES.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- MISCELLANEOUS: SECRETION OF SUBTILISIN IS ASSOCIATED WITH ONSET OF
CC SPOULATION, AND MANY MUTATIONS WHICH BLOCK SPOULATION AT EARLY
CC STAGES AFFECT EXPRESSION LEVELS OF SUBTILISIN. HOWEVER, SUBTILISIN
CC IS NOT NECESSARY FOR NORMAL SPOULATION.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D25319; BAA04989.1; -
DR EMBL: S51909; AAC60424.1; -
DR PIR: JH0778; JH0778.
DR HSP: P07518; IME.
DR MEROPS: S08.001; -
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.

```

```

DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Sporulation; Serine protease; zymogen; signal.
FT SIGNAL 1 .23
FT PROPEP 24 106 POTENTIAL.
FT CHAIN 107 381 SUBTILISIN NAT.
FT ACT_SITE 138 138 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 170 170 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 327 327 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT SEQUENCE 381 AA; 39507 MW; DAED4B16ED1BA092 CRC64;

Query Match 5.58; Score 188.5; DB 1; Length 381;
Best Local Similarly 22.58; Pred. No. 0.0084;
Matches 80; Conservative 61; Mismatches 131; Indels 83; Gaps 15;

QY 86 YTDSFVKNKTEENP-KKEDKYVYIAEPKDESGEKA--KELSLKNTKVLTYDRIENG 142
D 18 FTMKPSNMSQAAGKSSSTTEKRYIVGFQPMASAKKDIYSKGVQKQF-KYVMAA 76
QY 143 ALETTPNLDKIKQIEGSSVERAKQVPMNNAHKEIGVEAIDY-LKSINAP--FGKN 199
D 77 AATIDEKAVKELKKDPVAYVE-----DHIAHEVA--QSVPGISQIKAPALHSQG 126
QY 200 FDGRGVNISIDTGYTRHKAMRIDDDAKASMRKKEDLAKTDKNYMLSKIPAEVNYN 259
D 127 YTGSGNVKVAIVDIDSGIDSHPLNVRGGSF-----VPSETNRYQ 165
QY 260 GCKTVEYDDGRYFPHGWHINGILAGNTEODIKNFNGIDGIAPNAQISYKMYSDA 319
D 166 DG-----SSHGCHVAGTIA-----LNNSIGVLVAAPSASLYAKVLDST 205
QY 320 GSGFAGDETFHAIEDSIKHNVDVSVS-SGFTGLGEGEYMOAIRAKRAGIPMVVAT 378
D 206 GSG--QYSWIINGEIMAINSMNDYINMSLGPGTSLAKTYVDAAV-----SSGIVMAAA 259
QY 379 GNYATSSASSSMDLVANNHLKMTDGTGNTRTAHEDAIAV--ASAKNOTVEFDKY 431
D 260 GNESSGSGTST-----VGYPAKYPRTIAGVAVSSNQRASSFSSV 298

RESULT 13
SUBN_BACSU STANDARD; PRT; 645 AA.
AC P16396;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MINOR EXTRACELLULAR PROTEASE EPR PRECURSOR (EC 3.4.21.-).
GN EPR OR IPA-15R.
OS Bacillus subtilis.
OC Bacillus; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-168 / DB204;
RX MEDLINE-90340301; PubMed-2116590;
RA Brueckner R., Shoseyov O., Dol R.H.;
RT "Multiple active forms of a novel serine protease from Bacillus
RT subtilis.";
RL Mol. Gen. Genet. 221:486-490(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-89053875; PubMed-3142851;
RA Sioma A., Ally A., Ally D., Pero J.;
RT "Gene encoding a minor extracellular protease in Bacillus subtilis.";
RL J. Bacteriol. 170:5557-5563(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-168;
RX MEDLINE-95020537; PubMed-7934828;
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,

```

RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA Pressecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees."
RL Mol. Microbiol. 10:371-384(1993).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X53307; CA37392.1; -
DR EMBL; M22407; AAA22423.1; -
DR EMBL; X73124; CA51571.1; -
DR EMBL; 299123; CA51566.1; -
DR PIR; S11504; SUBSMP.
DR PIR; S39670; S39670.
DR HSSP; 099405; 1MPT.
DR Subtilisin; BG10561; epr.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILASIN.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Serine protease; zymogen; Signal; Complete proteome.
FT SIGNAL 1 27
FT PROPEP 28 103 POTENTIAL.
FT CHAIN 104 645 MINOR EXTRACELLULAR PROTEASE EPR.
FT ACT_SITE 142 142 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 172 172 CHANGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 326 326 CHANGE RELAY SYSTEM (BY SIMILARITY).
SQ SEQUENCE 645 AA; 69695 MW; 401AD5B60BE2EA CAC64;

Query Match 5.48; Score 187; DB 1; Length 645;
Best Local Similarity 22.1%; Pred. No. 0.02;
Matches 136; Conservative 74; Mismatches 191; Indels 214; Gaps 33;

OY 175 HARKIGVEEADIDYLSKSNAPKGFGRGMYI-SNIDTGTYYRH-KAMRIDDKAKSMR 232
DB 26 HAQNSSEKEVIVY-----KMKAGKETLDSADVEDQYKHLPAVAATVDETYKE 76
OY 233 FKE-----DLK-----GTD-----KNYV---LSDK-- 250
DB 77 LKQDDILYENNVSTFADSTDFKVLSDGDTSDNFQEMNLEPTQVQKAMKAGLTGKRI 136
OY 251 -----IPAFNYNGKATVE-----KYDGRDYFDPGHMIIAGILAGNDTEODIK 296
DB 137 KIAVIDSGISPHDLISLAGYSAVSYTSSYKDDNG-----HGTAVAGILGAK-----H 184
OY 297 NFNQIDGIAIPNAQITSYKMYSDAGSGFAGDETFHAIEDSKIHANVDVYSVSGFTGTLV 356
DB 185 NGYIGDIIAIPNAQITSYKMYSDAGSGFAGDETFHAIEDSKIHANVDVYSVSGFTGTLV 356
OY 357 GEKYQOAIIRALKAGIPVAVATGNAVTSASSSSMDLVANNHMKMDTGN---VTRTAHE 413
DB 243 ---LHDVAVKAYEGCVLLVAASGN-----DGKAKVAVPAAS 277
OY 414 DAIVASA--KNQTEFEDKVINIGSEFXY---RNIGAFEDKSKITTEDGTAKA-PSKLK 466
DB 278 SVAVASATNENKQNLASFST---GDEVEFSAAGTITSTYLNQYVATSGTSGQARPHAA 334
OY 467 FVYIGKQDDOLIGDLGKIAVMDRIYTKDKNAFKRAMDKGARAIVNVTVNYNRDN 526
DB 335 MFALLKQDPAPETNVOLEEMK-----KNIVDLGTA-----GRDO 369

OY 527 WTELPAWGVEAD-----EGTKSQVESISGDDGVKILNMI-----NPDKKT 566
DB 370 QREYGLIYKKAQTDATSAVAAEQAVKKAQDTKQI-----DLNKARELISQLPNSDAKT 423
OY 567 EVKR-----NKKEDFKDKL--EYQ---YVIDSEFSNK-PVNGDEKE----- 603
DB 424 ALHKRLDKVQSYRNVDKADKVAKEKRYTQGVDTAQTAINKLPGCTDKKNLQKRLDQV 483
OY 604 ---IDKFPAPD---TDKELYKEDI-----IYPAGS--TSMGPRIDLLKPDVSAPG 646
DB 484 KRYIAKQAKDKYAKAKESKKTVDVSQAQSAIGKLPPASSSEKTSLOKRLNKVSKTNL---- 539
OY 647 KNKSTLVNINCKST 661
DB 540 KTAQGSVSAAEKKT 554

RESULT 14
TAGB_DICDI
ID TAGB_DICDI STANDARD; PRT; 1905 AA.
AC P54683;
DT 01-OCT-1996 (rel. 34, Created)
DT 01-OCT-1996 (rel. 34, Last sequence update)
DT 20-AUG-2001 (rel. 40, Last annotation update)
DE PRESTALK-SPECIFIC PROTEIN TAGB PRECURSOR (EC 3.4.21.-).
GN TAGB.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Eukaryota; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=95262903; PubMed=7744252;
RA Shaulsky G., Kuspa A., Loomis W.F.;
RT "A multidrug resistance transporter/serine protease gene is required
RT for prestalk specialization in Dictyostellium."
RL Genes Dev. 9:1111-1122(1995).
CC -1- FUNCTION: INTERCELLULAR COMMUNICATION VIA TAGB MAY MEDIATE
CC INTEGRATION OF CELLULAR DIFFERENTIATION WITH MORPHOGENESIS.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO PEPTIDASE FAMILY
CC S8; ALSO KNOWN AS THE SUBTILASE FAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION, BELONGS TO THE ATP-BINDING
CC TRANSPORT PROTEIN FAMILY (ABC TRANSPORTERS), MDR SUBFAMILY.
CC -1- SIMILARITY: STRONG, TO TAGC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U20432; AAA62212.1; -
DR HSSP; P13569; 1NBD.
DR Dictydb; DD02059; tagb.
DR InterPro; IPR003593; AAA.
DR InterPro; IPR003439; ABC_transporter.
DR InterPro; IPR001140; ABC_transporter_tmam.
DR InterPro; IPR001687; ATP_GTP_A.
DR InterPro; IPR000209; Peptidase_S8.
DR Pfam; PF00664; ABC_membrane; 1.
DR Pfam; PF00005; ABC_tran; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00136; SUBTILASE_ASP; FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Hydrolyase; Serine protease; ATP-binding; Transport; Transmembrane;
KW SIGNAL 1 ? POTENTIAL.

SUBTILASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D00264; BAA00186.1; -
DR PIR: A00971; SUBSS.
DR PIR: A41448; A41448.
DR HSSP: P07518; IMEE.
DR MEROPS: S08.001; -
DR InterPro: IPR000209; Peptidase_S8.
DR Pfam: PF00082; Peptidase_S8; 1.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Hydrolyase; Sporulation; Serine protease; Zymogen; Signal.
FT SIGNAL 1 30
FT PROPEP 31 106
FT CHAIN 107 381
FT ACT_SITE 138 138
FT ACT_SITE 170 170
FT ACT_SITE 327 327
FT CONFLICT 191 191
FT CONFLICT 365 365
SQ SEQUENCE 381 AA; 39467 MW; 2251BADE22B4824F CRC64;
POTENTIAL.
SUBTILISIN AMYLOSACCHARITICUS.
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
CHARGE RELAY SYSTEM (BY SIMILARITY).
S -> A (IN REF. 2).
N -> D (IN REF. 2).

Query Match

5.4%; Score 185.5; DB 1; Length 381;
Best Local Similarity 22.8%; Pred. No. 0.012;
Matches 81; Conservative 59; Mismatches 126; Indels 89; Gaps 16;

QY 86 YTDSEFVNKNTENP-KKEDKVYVIAEFKDESEGERAI--KELSLKNTKVLTYDRIFNGS 142
DB 18 FTMAFSNMSAQAAGKSSTEKKYIVGFKOTMSAMSAKKKDYISEKGKVKOKF-KYVNAA 76
QY 143 ALETPDNLKIKOIEGISSVERAKVOPMMNHAKKEIGVEAIDY-LKSINAP--FGKN 199
DB 77 AATLDEKAVKELKKDPVAAYVEE-----DHIAHEYA--QSVPYGISQIKAPALHSQG 126
QY 200 FDGGMVINSNDTGTDRYHKKAMRIDDAKASMRKKEDLKCTDKNYWLSDKIPIAFENYNN 259
DB 127 YTGSNVAVYIDSGIDSHPLNVRGGSF-----VPSETNPYQ 165
QY 260 CGKITVEKYDDGRDYFDPGHMHIAGIILAGNDTEODIKNFNGIDGIAPNAOIFSYKMYSDA 319
DB 166 DG-----SSHGTHVAGTIAA-----LNNSIGLVGSPSASLYAKVLDST 205
QY 320 GSGFAGDETFMHAIEDSIKHNVDVSVS---SGFTGTGLVGEKRYWQAIRALKRAGIPMV 375
DB 206 GSG--QYSMIINGIEMAISSNMMDIYNNISLGGPSGSLTKTYVDK--AV---SSGIYVA 256
QY 376 VATGNVATSASSSSMDLVANNHLMKMTDGNVTRTAHEDAIAV--ASAKNOTVEE 428
DB 257 AAAGNEGSSGSSST-----VGYPAKYPTSTIAGAVNSSNORASF 295

Search completed: December 7, 2001, 11:49:51
Job time: 205 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 11:46:10 ; Search time 28.64 Seconds
(without alignments)
3391.230 Million cell updates/sec

Title: US-09-590-991-6
Sequence: 1 YPVVLADTSSSEDALNISDK.....PEKNIKSTLVNKGKSTYGV 664

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP_invertebrate: *
6: SP_mammal: *
7: SP_mhc: *
8: SP_organelle: *
9: SP_phage: *
10: SP_plant: *
11: SP_rodent: *
12: SP_virus: *
13: SP_vertebrate: *
14: SP_unclassified: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3454	100.0	2119	2 Q9AHT5	Q9AHT5 streptococc
2	3342.5	96.8	2144	2 Q9S4M8	Q9S4M8 streptococc
3	453	13.1	1946	2 Q48545	Q48545 lactobacill
4	438.5	12.7	1902	2 Q9A102	Q9A102 lactococcu
5	424.5	12.3	1849	2 Q9SAK2	Q9SAK2 lactobacill
6	410.5	11.9	1585	2 Q9F8Q4	Q9F8Q4 streptococc
7	361	10.5	1647	2 Q9A180	Q9A180 streptococc
8	340	9.8	692	2 Q9EXK0	Q9EXK0 bacillus ps
9	298.5	8.6	757	2 Q9K6G6	Q9K6G6 bacillus ps
10	286.5	8.3	1150	2 Q53637	Q53637 streptococc
11	285.5	8.3	799	2 Q9KEM1	Q9KEM1 bacillus ha
12	279.5	8.1	1134	2 Q9L508	Q9L508 streptococc
13	257.5	7.5	1130	2 Q9K2K8	Q9K2K8 streptococc
14	256	7.4	731	10 Q39547	Q39547 cucumis mel
15	252.5	7.3	840	10 Q9F112	Q9F112 arabisopsis
16	251.5	7.3	1433	2 Q45616	Q45616 bacillus su
17	250	7.2	801	10 P93221	P93221 lycopersico
18	246	7.1	746	10 Q9LMA3	Q9LMA3 lycopersico
19	242	7.0	754	10 Q65834	Q65834 lycopersico

20	241	7.0	809	10 Q9SPA0	Q9SPA0 lycopersico
21	239.5	6.9	747	10 Q82007	Q82007 lycopersico
22	238.5	6.9	745	10 Q96478	Q96478 lycopersico
23	238	6.9	813	10 Q43546	Q43546 lilium long
24	234	6.8	736	10 Q9FYC7	Q9FYC7 oryza sativ
25	232.5	6.7	743	10 Q9LMA4	Q9LMA4 lycopersico
26	232	6.7	856	10 Q9SUN6	Q9SUN6 arabidopsi
27	231.5	6.7	816	10 Q9S4V5	Q9S4V5 arabidopsi
28	229	6.6	713	10 Q9F1M5	Q9F1M5 arabidopsi
29	228	6.6	747	10 Q65836	Q65836 lycopersico
30	227	6.6	754	10 Q65835	Q65835 lycopersico
31	223	6.5	747	10 Q9SAN2	Q9SAN2 lycopersico
32	220.5	6.4	666	10 Q9ZK46	Q9ZK46 lycopersico
33	219.5	6.4	766	10 Q9LZS6	Q9LZS6 arabidopsi
34	219.5	6.4	815	10 Q64481	Q64481 arabidopsi
35	212.5	6.2	749	10 Q9L1L8	Q9L1L8 arabidopsi
36	211.5	6.1	761	10 Q82777	Q82777 lycopersico
37	208	6.0	783	10 Q04190	Q04190 arabidopsi
38	206.5	6.0	697	10 Q9FTG1	Q9FTG1 arabidopsi
39	206	6.0	779	10 Q9LNU0	Q9LNU0 arabidopsi
40	204.5	5.9	741	10 Q9FCU3	Q9FCU3 arabidopsi
41	203.5	5.9	777	10 Q9ZS44	Q9ZS44 lycopersico
42	202.5	5.9	693	10 Q9F1M8	Q9F1M8 arabidopsi
43	202	5.8	745	10 Q04678	Q04678 lycopersico
44	198.5	5.7	707	10 Q9FEU4	Q9FEU4 arabidopsi
45	196	5.7	706	10 Q81324	Q81324 arabidopsi

ALIGNMENTS

```

RESULT 1
ID Q9AHT5 PRELIMINARY; PRT; 2119 AA.
AC Q9AHT5;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SERINE PROTEASE (FRAGMENT).
GN PRTA.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M;
RX MEDLINE=2116976; Pubmed=11179332;
RA Witzmann T.M., Heinrichs J.H., Adamou J.E., Erwin A.L., Kunsch C.,
RA Choi G.H., Barash S.C., Rosen C.A., Masure H.R., Tuomanen E.,
RA Gayle A., Brewah Y.A., Walsh W., Barron P., Lathigra R., Hanson M.,
RA Langermann S., Johnson S., Koenig S.;
RT "Use of a whole Genome Approach To Identify Vaccine Molecules
RT Affording Protection Against Streptococcus pneumoniae Infection.";
RT Infect. Immun. 69:1593-1598(2001).
DR EMBL; AF291699; AAK19159.1; -.
KW Protease.
FT NON_TPR.
SQ SEQUENCE 1 1
2119 AA; 238226 MW; 517F9B7FB960A6A CRC64;

```

Query Match 100.0%; Score 3454; DB 2; Length 2119;

Best Local Similarity 100.0%; Pred. No. 5.2e-174; Matches 664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 YPVVLADTSSSEDALNISDKKEVAENKEHENIHSAHMETSDQFEKKTAIVIKKEVSKN 60
DB 1 YPVVLADTSSSEDALNISDKKEVAENKEHENIHSAHMETSDQFEKKTAIVIKKEVSKN 60
QY 61 PYVDNNNTSNEAKTKENSNSGQDYTDSPFNKNTENPKKEDKVVYIAEFQDKSGEKAI 120
DB 61 PYVDNNNTSNEAKTKENSNSGQDYTDSPFNKNTENPKKEDKVVYIAEFQDKSGEKAI 120

```

QY 121 KELSLKNTKVLTYVDRIJFNGSAIETTPDNLKIKQIEGISSVERAQKQVPMNNHARKEI 180
 DB 121 KELSLKNTKVLTYVDRIJFNGSAIETTPDNLKIKQIEGISSVERAQKQVPMNNHARKEI 180
 QY 161 GVEEADIVLKSINAPFGKFDGRGVISNIDTGTDRHAKMRIDDADAKASMRKKEDLKG 240
 DB 161 GVEEADIVLKSINAPFGKFDGRGVISNIDTGTDRHAKMRIDDADAKASMRKKEDLKG 240
 QY 241 TDKNWLSDKIPHAENYNGKITVEKYDDGRDYPDPHGMHAGIAGNDTEEDIKNFNG 300
 DB 241 TDKNWLSDKIPHAENYNGKITVEKYDDGRDYPDPHGMHAGIAGNDTEEDIKNFNG 300
 QY 301 IDGIAPNAQIFSYKMYSDAGSGFAGDETFHAIEDSIKHNVVSVSSGFTGTGLVGEKY 360
 DB 301 IDGIAPNAQIFSYKMYSDAGSGFAGDETFHAIEDSIKHNVVSVSSGFTGTGLVGEKY 360
 QY 361 WOATLALRKAGIPMYVATGNVATSSSSWDLVANNHLMKMTDGTGNTRTAAHEDALAVAS 420
 DB 361 WOATLALRKAGIPMYVATGNVATSSSSWDLVANNHLMKMTDGTGNTRTAAHEDALAVAS 420
 QY 421 AKNOTVEPKVNGESPFYRNIGAFEDKSKITTTNEDGTAKPSKLEFYIIGKQDDDLG 480
 DB 421 AKNOTVEPKVNGESPFYRNIGAFEDKSKITTTNEDGTAKPSKLEFYIIGKQDDDLG 480
 QY 481 LDLRKLIAMDRITTKDLKNAEKKAMDKGARAIMVNTVNYNRDMWTELPAMGYEADG 540
 DB 481 LDLRKLIAMDRITTKDLKNAEKKAMDKGARAIMVNTVNYNRDMWTELPAMGYEADG 540
 QY 541 TKSQVFSISGDDGVKLMNINPDKTEYVRNKKEDFKDLBOYYPIDMESFNKNRNGVD 600
 DB 541 TKSQVFSISGDDGVKLMNINPDKTEYVRNKKEDFKDLBOYYPIDMESFNKNRNGVD 600
 QY 601 EKEIDFKAPDPTDKELYKEDIIVPAGSTSMGPRIDLLKPDVAPKKNKSTLVNNGKS 660
 DB 601 EKEIDFKAPDPTDKELYKEDIIVPAGSTSMGPRIDLLKPDVAPKKNKSTLVNNGKS 660
 QY 661 TYGY 664
 DB 661 TYGY 664

RESULT 2
 Q9S4M8 PRELIMINARY: PRT; 2144 AA.
 AC Q9S4M8:
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE CELL WALL-ASSOCIATED SERINE PROTEINASE PRECURSOR PRTA.
 GN PRTA.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OK NCBI_TaxID=1313;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=3.B;
 RA Beche G., ten Thoren E., Bongaele R.J.M., Heinz H.-P., Zysk G.;
 RT "Cloning and sequencing of a novel surface protease of Streptococcus
 pneumoniae".
 RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF17143; AAD48399.1; -.
 DR HSSP; P07518; IMEE.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR003137; PA.
 DR InterPro: IPR000209; Peptidase_S8.
 DR InterPro: IPR001680; WD40.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 2.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
 DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.

DR PROSITE; PS00138; SUBTILASE_SER; UNKNOWN_1.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
 KM Signal.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 2144 CELL WALL-ASSOCIATED SEI
 FT PRTA.
 SQ SEQUENCE 2144 AA; 240725 MW; 2052511470741331 CR...

Query Match 96.8%; Score 3342.5; DA 2; Length 2144;
 Best Local Similarity 96.3%; Pred. No. 4.1e-168;
 Matches 642; Conservative 14; Mismatches 6; Indels 5; Gaps 1;

QY 3 VVLADTSSSEDALISDKREKY-----AENKEHENHISAMETSDPFKEKTAIVIKEEV 57
 DB 24 VVLADTSSSEDALISDKREKYVDKTEENKHHIDHAIETSDTEKKTIIIEKEEV 83
 QY 58 SKNPVIDNNTSNEBAKIKENSNSKSGDYDSEFNKNTENPKKEDKVVYIAEFKDESGE 117
 DB 84 SKNPVIDTKTSNEBAKIKENSNSKSGDYDSEFNKNTENPKKEDKVVYIAEFKDESGE 143
 QY 118 KATKELSLKNTKVLTYVDRIJFNGSAIETTPDNLKIKQIEGISSVERAQKQVPMNNHAR 177
 DB 144 KATKELSLKNTKVLTYVDRIJFNGSAIETTPDNLKIKQIEGISSVERAQKQVPMNNHAR 203
 QY 178 KEIGVEEADIVLKSINAPFGKFDGRGVISNIDTGTDRHAKMRIDDADAKASMRKKED 237
 DB 204 KEIGVEEADIVLKSINAPFGKFDGRGVISNIDTGTDRHAKMRIDDADAKASMRKKED 263
 QY 238 LKGTDKRWYLSKIPHAENYNGKITVEKYDDGRDYPDPHGMHAGIAGNDTEEDIKN 297
 DB 264 LKGTDKRWYLSKIPHAENYNGKITVEKYDDGRDYPDPHGMHAGIAGNDTEEDIKN 323
 QY 298 FNGIDGIAPNAQIFSYKMYSDAGSGFAGDETFHAIEDSIKHNVVSVSSGFTGTGLVG 357
 DB 324 FNGIDGIAPNAQIFSYKMYSDAGSGFAGDETFHAIEDSIKHNVVSVSSGFTGTGLVG 383
 QY 358 EKYWOATLALRKAGIPMYVATGNVATSSSSWDLVANNHLMKMTDGTGNTRTAAHEDAL 417
 DB 384 EKYWOATLALRKAGIPMYVATGNVATSSSSWDLVANNHLMKMTDGTGNTRTAAHEDAL 443
 QY 418 VASAKNOTVEPKVNGESPFYRNIGAFEDKSKITTTNEDGTAKPSKLEFYIIGKQDD 477
 DB 444 VASAKNOTVEPKVNGESPFYRNIGAFEDKSKITTTNEDGTAKPSKLEFYIIGKQDD 503
 QY 478 LIGDLRGKTIAMDRITTKDLKNAEKKAMDKGARAIMVNTVNYNRDMWTELPAMGYEA 537
 DB 504 LIGDLRGKTIAMDRITTKDLKNAEKKAMDKGARAIMVNTVNYNRDMWTELPAMGYEA 563
 QY 538 DEGTKSQVFSISGDDGVKLMNINPDKTEYVRNKKEDFKDLBOYYPIDMESFNKNRPN 597
 DB 564 DEGTKSQVFSISGDDGVKLMNINPDKTEYVRNKKEDFKDLBOYYPIDMESFNKNRPN 623
 QY 598 VGDEKETIDFKAPDPTDKELYKEDIIVPAGSTSMGPRIDLLKPDVAPKKNKSTLVNIN 657
 DB 624 VGDEKETIDFKAPDPTDKELYKEDIIVPAGSTSMGPRIDLLKPDVAPKKNKSTLVNIN 683
 QY 658 GKSTYGY 664
 DB 684 GKSTYGY 690

RESULT 3
 Q48545 PRELIMINARY: PRT; 1946 AA.
 AC Q48545:
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE PROTEINASE (PRTB) PRECURSOR (PRTB).
 GN PRTB.
 OS Lactobacillus delbrueckii.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;

OC Lactobacillus.
 OX NCBI_TaxID=1584;
 RN [1]
 RX SEQUENCE FROM N.A.
 RA MEDLINE=6236017; PubMed=8655480;
 RA Gilbert C., Atlan D., Blanc B., Portallier R., Germond J.E.,
 RA Lapierre L., Mollet B.;
 RT "A new cell surface proteinase: sequencing and analysis of the *prtB*
 RT gene from *Lactobacillus delbrueckii* subsp. *bulgaricus*.";
 RL J. Bacteriol. 178:3059-3065(1996).
 DR EMBL, LA8487; AAC41529.1; .
 DR HSSP; P00782; 1SU6.
 DR InterPro; IPR003137; PA.
 DR InterPro; IPR00209; Peptidase_S8.
 DR Pfam; PF00225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 3.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; UNKNOWN_1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 DR Signal.
 RN CHAIN 1 34 POTENTIAL.
 RN SIGNAL 193 1946 PROTEINASE.
 SQ SEQUENCE 1946 AA; 212315 MW; 21EF17D02E79C6A0 CRC64;

Query Match 13.1%; Score 453; DB 2; Length 1946;

Best Local Similarity 25.7%; Pred. No. 1.6e-15;
 Matches 175; Conservative 114; Mismatches 261; Indels 130; Gaps 28;

OY 5 LADTSSSEDALNIDSEKVAENKEHENISAMETSDQFEKKT-----AVIKKEVYVK 59
 DB 27 LAGTFOGSAFAFOAASOETFEPRASRAALTYLQEOGRYNAKKSKYKFEAKEDQROASG 86
 OY 60 NPVIDNTSNEAKIKENSNSKSG-DYTSFVNKNTEPNKEDKVVYIAEFKESGEK 118
 DB 87 QAVSKKNESSRVLY---SLKNSAFDHT---SKPTGSAASVKKIEQASD-QVKDGOEK 138
 OY 119 AIKELSLKNTKVLTYTDFNGSAIETTPDNLKIKQIEGSSVERAKQVPMNNHARK 178
 DB 139 YIKOVEITGNKVRQGYLVNAFSDIMDDIDKVKDLPQVKNVTTPKVVHP----- 191
 OY 179 EIGVEEADLYLKSINAPP-GNEDGRGVTSNIDGTGYRHKAMRIDDD-----AKASMR 232
 DB 192 ---DDEASDAQVQDWQOEKLGEGMVISIIDTGIDSSHQDLGSGVSTALSKSEVE 248
 OY 233 FKREDLGTGKNWLSKIPAFENYNGKTTVEKYDDGRYFPRGHMIGIAGNTE 292
 DB 249 SDKSKL-GHGKTY--TEKVPYGVYADKNQIV--DNGCG--EMHGQHVAGIAGAN-- 297
 OY 293 QDIKNFNGIDGIAPNAOIFYKMTS-DAGSGFAGDETFPHAIEDSIKHNVDVSVSGFT 351
 DB 298 -----GVKGVAPDADQLAMKVFSSNAKNSGAYDDDLISIEDSVKIGAVINMSLGSV 351
 OY 352 GTGL-VEKTYWQAIRALKAGIPVAVATGNVATSSSSSMDLVAN--NHLKMTDTGNVT 407
 DB 352 SSDGSPDPOQAAVAKASAGVIVINISAGNSGVASTADGNPVNNTGSELSTVGTPCVT 411
 OY 408 RTAHEDAIIVASAKNOTVERDKY--NIGSESEKRYN--IAAFDPSKITTEGCTKA-- 461
 DB 412 ----PALTVYASAEKNSVTTDTVKDELGVTFSSNELKA---AQVTTQLSNNVSVL 462
 OY 462 PSKLKPYVIGQDQDLIG--LDLRGIAMVDR-IYTKDLNAFKKAMDGARAIVNVN 517
 DB 463 TKRKLKLVDMGAGADTYAEKKAKEVGQGLAVYKRGATYFSAKAVANAKA--AGAAGIYIYN 520
 OY 518 TVNYNRDNMTTELPAMGVEADEGTSQVFSISGDGVKLYMNMIPDKKTEVRKNNKDEK 577
 DB 521 -----SEDDG--LLSMSLDDKTFPTLGSKADGK 547
 OY 578 DKLEOYYPIDMESFNNSKNPNYGD---KEIDFKAPDIDKLYKEDIIVPAGSTSMGPRI 634
 DB 548 FWLQOKKVRASRLKFTGALTIDNSRAGKMSDF-----TSMGPTP 586

OY 635 DLLKPDVSAAPGNKIKSTLN 654
 DB 587 ELDFKPEITAPGKXIYSLAN 606

RESULT 4

ID 09A102 PRELIMINARY; PRT; 1902 AA.

AC 09A102;

DT 01-JUN-2001 (Tremblrel, 17, Created)

DT 01-JUN-2001 (Tremblrel, 17, Last sequence update)

DT 01-JUN-2001 (Tremblrel, 17, Last annotation update)

DE PRP PRECURSOR.

GN PRP.

OS Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).

OC Plasmid PHP003.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

OC Lactococcus.

OX NCBI_TaxID=1359;

RN [1]

RP SEQUENCE FROM N.A.

RC SPRAIN-HP;

RA Christensen C., Pillidge C.J., Ward L.J.H., O'Toole P.W.;

RT "Characterization of a lactococcal plasmid in *Lactococcus lactis* subsp.

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF247159; AAK27981.1; .

KW Signal; Plasmid.

FT SIGNAL 1 33 POTENTIAL.

SQ SEQUENCE 1902 AA; 200497 MW; 50B3D9DAE575B2B CRC64;

Query Match 12.7%; Score 438.5; DB 2; Length 1902;

Best Local Similarity 27.3%; Pred. No. 9.1e-15;
 Matches 166; Conservative 91; Mismatches 229; Indels 111; Gaps 25;

OY 78 NSNKSQGDYDTSFVNKNTEPNKED--KVYIAEFKDESEKAIKELSLKNTKVLTYT 134
 DB 84 NKVQOQDIYVDVYOMSA--PASENGTLRTDYSSFAIEQETNKVIAAASVKAVEGVT 142
 OY 135 -----YDRFNSAITTPDNLKIKQIEGSSVERAKQVPMNNHARKIEIGVEEAD 187
 DB 143 QOTAGESGVYVNFSTKVRVVDIPKQIAGVKTIVTLAKYYPDADAKANSMAVQAVWS 202
 OY 188 YLKSINAPFGKNGRGVVISNIDGTGYRHKAMRIDDAKASMFKEEDL-KGTD--K 243
 DB 203 NYK-----YKGEGVSVYIDSGIDPTHKDKRLSD--KDKYLTSDVEKFTDYKHN 251
 OY 244 NYWLSDKIPRAFNYY-NGKLTVEKYDDGRYFPRGHMIGIAGNTEQDIKNFNGID 302
 DB 252 GRYNSKVPYGFENVADNNDITTDKVD-----EGHGMHVGIIIGANGTGD--PAKSVY 303
 OY 303 GIAPNAQIFSTKMTSDAC-SGFAGDETFPHAIEDSIKHNVDVSVSGF-GTGLVGEKY 360
 DB 304 GVAPEAQILAMKVFVNSDTSTGSDTLVSAIEDSAKIGAVULNMSLSDSGNQTLEDP 363
 OY 361 WQAIRALKKAGIPVAVATGNVATSSSSSMDLVANNNHLMKMTDTGNVTAAHEDAIIVAS 420
 DB 364 IAAVQVANNESGTAIVISAGNSGTSGSAT--EGVKNQDIYGLDDNEVNGVPGTSKGAITYAAS 421
 OY 421 AKNOTVERDKVNI-----GGESEFKY--RNIGAFDPKSKITTEGCTKAPSKLIFYI 470
 DB 422 AENDVITQAVTITIDGTGLGPEITQLSSNDPFGSFQKKFYVVKDAISGNLSK----- 475
 OY 471 GKQDQDLIGDLNGKTIAMVDR-IYTKDLNAFKKAMDGARAIVNVTVNYNRDNMTTE 529
 DB 476 GKVADYT--ADAKGKIAIVRGELTFEDDKQKYQA--AGAAGLIYN-----NDGTATP 525
 OY 530 LPAMGVEADEGTSQVFSISGDGVKLYMNMIPDKKTEVRKNNKEDFKDLQGYEYID 587
 DB 526 VTSMALT--TTPPTFELSSVTGOKLVQVYTAHPDSSLGK----- 563
 OY 588 MESFNSKNPNVDEKIDFKAPDIDKLYKEDIIVPAGSTSMGPRI DLLKPDVSAAPGK 647

Db	564	-----	IATLVPN	---	QAVTDDKM	--	SDFTSGVPSNLSFKPDITAPG	602
Oy	648	NKSTLN	654					
Db	603	NIMSTON	609					
RESULT	5							
09SAK2	ID	09SAK2	PRELIMINARY:	PRT:	1849	AA.		
AC	09SAK2							
DT	01-MAY-2000	(TREMBLrel. 13, Created)						
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)						
DT	01-JUN-2001	(TREMBLrel. 17, Last annotation update)						
DE	CELL ENVELOPE-ASSOCIATED PROTEINASE.							
GN	PRH.							
OS	Lactobacillus helveticus.							
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;							
NCBI	Lactobacillus.							
NCBI	_taxid=1587;							
NP	[1]							
RP	SEQUENCE FROM N.A.							
RC	STRAIN=CNR232.							
RX	MEDLINE=99350424; PubMed=10419958;							
RA	Pederson J.A., Mileski G.J., Welmer B.C., Steele J.L.;							
RT	"Genetic characterization of a cell envelope-associated proteinase							
RT	from Lactobacillus helveticus CNR32.";							
RL	J. Bacteriol. 181:4592-4597(1999).							
DR	EMBL: AF133727; AAD50643.1; .							
DR	HSSP: P00782; 1SUC.							
DR	InterPro: IPR001444; Flag_bf-rod.							
DR	InterPro: IPR003137; PA.							
DR	InterPro: IPR002029; Peptidase_s8.							
DR	Pfam: PF02225; PA.1.							
DR	Pfam: PF00082; Peptidase_s8; 3.							
DR	PRINTS: PR00723; SUBTILISIN.							
DR	PROSITE: PS00588; FLAGELLA_BB_ROD; UNKNOWN_1.							
DR	PROSITE: PS00136; SUBTILASE_AS; UNKNOWN_1.							
DR	PROSITE: PS00137; SUBTILASE_HIS; 1.							
DR	PROSITE: PS00138; SUBTILASE_SER; 1.							
DR	SEQUENCE 1849 AA; 199614 MW; 219F0D44B15A091F CRC64;							
SO								
Query Match	12.3%	Score 424.5;	DB 2;	Length 1849;				
Best Local Similarity	24.1%;	Pred. No. 4.8e-14;						
Matches 160;	Conservative 113;	Mismatches 258;	Indels 133;	Gaps				
Oy	31	ENHSAMETSOOFKREKTAIVIKEKEVSKNPYIDNNTSNEEAKIKIEENSNSQGDYDTSF	90					
Db	27	QOVKASVDSQFTVEKSKTKRAESTTANLTNKAVEAQLAKGVPFHLITVNOQDYVDVI	86					
Oy	91	V-----NKNTENPKKEDKVYIAEFKKDESGEKAIKE-ISSLKNTKULY	133					
Db	87	VQLSATPATNGSVANSSSSAEIDEASKVI-----ANQASIKEKVKAITDQAIK	137					
Oy	134	TYDRFNGSAIETTPDNLKIKOIGISIVEAKQVKYPMANNAKREIGEEALIDYKSTN	193					
Db	138	SYGVYVNGFATRAKKKIDOKLRNITGVASVTLAK-----VYIAN-----DSSADMANAVS	187					
Oy	194	APFGK-NFDGRCGVISNIDTGTDRHKKAMRIDDAKASMRFEKEDLKGTDKNY--WLSDK	250					
Db	188	TYVMNNYKKYKGEETVYSIIDGTIDPHKKDLRLSDSKVKLTCKDVNAFTRESGYGRFTDK	247					
Oy	251	IPHANVY-NGKRIVEKYIDGCRDYFDRHGHIAGILLAGNPTDEDIKNPNNGIDGIAPNAQ	309					
Db	248	VPGYGNYSDDNDNLT---DDNPS--EEOHGMHVAIGIVANGTADSV---NSVYGAAPBAEQ	298					
Oy	310	IFSYMYSDAGSGFAGDET-MFHAIEDSIKHNVDVAVSVSSG-FTGAGLGEKYYMOAIRAL	367					
Db	299	LLAKAFENSDSSASTDSTSIIGALIDDSAKLCAADVPLNLSGVSSEGTDEDPYAAVERA	358					
Oy	368	RKAGIPVAVATGVNTVATSSSSSWDLVANNHLKMTDTGVNTRTAAHEDALVAASAKNQVTE	427					

```

Db      359  TKGTAAVYLSAGNSGR--SNSELEGVNRKYYGNDPDEITGNGPTARSATTAASAEKTRAT 416
Qy      428  EDKVNIGESFYKRYRNIG-----AFDFKSKITTNEDGTAPSKLAKVYIGKGOD 475
Db      417  TGGVITTSADGRTTTIAGPEATOLSEGTDRAPFNDDKFFVVKD-----KNGNLGTGSA 468
Qy      476  ODLIGLDLGLKAVMDR---ITYTKDLKNKFRKAMOKGARAIVWTVNYNNDNTTELA 532
Db      469  KQYTS- VNGKIALYKRGELTFT---DKQKTAQEGAGAGLIIVN-----NKAGDING 516
Qy      533  MGEYADEGTRKSOVFSISGDDGVKLMNMT--NPDKTEYKRNKKEFKDLQYYPIDMES 590
Db      517  MLTNGGFPTA---GLSATSGEKLKYYVAHDEALVKY-----IYVQA 556
Qy      591  FUSNKPWVDEKKEIDFKFAPDPTDKELKREDIIVPAGSTGWSGRIDLTLKPDVASAPGNKX 650
Db      557  LNNSS-----ARQTD-----LMSDFSTYSGPTSSLAEPDISAPGHIW 593
Qy      651  STLN 654
Db      594  STON 597

RESULT 6
Q9F8Q4 PRELIMINARY; PRT; 1585 AA.
ID      Q9F8Q4
AC      Q9F8Q4
DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      CELL ENVELOPE PROTEINASE.
GN      PRTS.
OS      Streptococcus thermophilus.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC      Streptococcus.
OX      NCBI_Taxid=1308;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20509797; PubMed=11055922;
RA      Fernandez-Espia M.D., Garault P., Monnet V., Rul F.;
RT      "Streptococcus thermophilus Cell Wall-Anchored Proteinase: Release,
RT      Purification, and Biochemical and Genetic Characterization.";
RL      Appl. Environ. Microbiol. 66:4772-4778(2000).
DR      EMBL, AF243528; AAC09771.1.-.
DR      InterPro: IPR001899; Gram_pos_anchor.
DR      InterPro: IPR003137; PA.
DR      InterPro: IPR000209; Peptidase_S8.
DR      Pfam: PF00746; Gram_pos_anchor; 1.
DR      Pfam: PF02225; PA; 1.
DR      Pfam: PF00082; Peptidase_S8; 3.
DR      PRINTS: PR00723; SUBTILISIN.
KW      Envelope protein.
SQ      SEQUENCE 1585 AA; 169873 MW; AD210BCELF9A399 CRC64;

Query Match      11.9%; Score 410.5; DB 2; Length 1585;
Best Local Similarity 24.3%; Pred. No.2.1e-13
Matches 160; Conservative 104; Mismatches 245; Indels 149; Gaps 24;
Qy      44  KEKTAVAIVKEKVVSKNPYI-----DNNTSNEBAKTEKENSNSKSGDYDTS 89
Db      2  KKEKFFSLRKYYIGIVSVSLGAVFLFAGAPSVAADELISLVETKY-----EATVPDA 53
Qy      90  FVNKA-TEPKKEDKVVYIAE-----FKDKESEBAKTELKSLKNKRVLYTYDRIFNG 141
Db      54  IYSESASESPVVEELVYDVSVEATSYDVYTTTDEEBTPGSEALENSANTEVETQ----- 107
Qy      142  SAIEETPDMLDKIKQIEGSISSVERAQKQVPMNHNHAKKEIGVEALDYLSINAP--FGKN 199
Db      108  PAVETPAISSEKKVVEEEKISVADETTAL-----TNOEBAKQONIDSNTIITVPKAVYISG 161
Qy      200  FDGRGMVSNIDTGYDTRHKAIRIDDKAKSMRFKE-----DLKGTCKNWVLSDKIPHA 254

```


DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF02225; PA: 1.
 DR Pfam: PF00082; Peptidase_S8; 2.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
 KM Hydrolyase.
 FT NON_TER 1
 SEQUENCE 692 AA; 75155 MW; E53D6CB38E6B6D9 CRC64;

Query Match 9.8%; Score 340; DB 2; Length 692;
 Best Local Similarity 24.9%; Pred. No. 3.3e-10;
 Matches 144; Conservative 77; Mismatches 150; Indels 208; Gaps 26;

QY 126 LKNTKLYLYDRIFNGSAIETTPDNLKIKQIEGSSVERAOKVOPMMNHAKKEIGVEEA 185
 DB 1 LPSSKVRKKFNLVNGFSELEKGDLDVLKQV---SVERVQV-----AQYEANLDGS 51
 QY 186 I-----DYLKSTINADPGKNGFGKGVNISNIDTGTDRHKAHRIDDAKASMRKKEDLKG 240
 DB 52 VPFIGDHYRGWFSRSDGELTGKGVKAVIDTIGIDYRH-----PDLQA 94
 QY 241 TDKNWMLSDKIPRANVNYNGCKITVEKYDDGDRYDP-----HGHIAGILAG 288
 DB 95 N-----YKGGDYVD-YD-----HDPMETQSTQGVPTLHGTHVAGITAA 132
 QY 289 NDTEODIKNFNGIDGIAFNAQIFSYKMYSDAGSGFAGDETFMHAIEDSIKHNVDVSVSS 348
 DB 133 N-----GGKGVAPAEADIAIRALGPGGQGT--EQVIEALEKAVADGVDLNLSL 181
 QY 349 GFTGTGLVGEKYMQAIRALKRAGIPVAVATGNYATSSASSWDLVANNHLMKMTDTGNVTR 408
 DB 182 GTVVG---PDMPTSVALLDK----- 198
 QY 409 TAAHDAIAVASAKQTEVEFOKVNIGSEFKYRNIG-----FEKSKITINE 456
 DB 139 -AVEGVAIVTNSNGSGPMMTVSGSPSTKAISVGASAPRIKTPYATVFGDKELTLP 257
 QY 457 DGTAKPSKLK---RVYIYGKGDODLIGDLRGKIAVMDR--IYTKDLKNAFKKAMDG 509
 DB 258 MCGTAPMALKRDFPLVDGGLGTEEDLDELEVGSVLVAKRGILPTPEKAHNA-KKA---G 313
 QY 510 ARAIVAVTVNTYRNNDNMTLP-AMGYEADGTSQVFSISGDDGVKLMNINPDKTEV 568
 DB 314 AKAMIIYN-----NLPGAFVGYEGVKLPVVSVTKEG----- 347
 QY 569 KRNNEDEKDKLEQYYPIOMESFNKPNVDEKEIDFKFADDTKELYK--EDIIVAG 626
 DB 348 -----EYLDKLEAH-----KEL-----TIRTIYREDEDTAPPS 378
 QY 627 STSMGPRIIDL-LKPDVAPGKNIKSTLN---VINGKS 660
 DB 379 SR-GPVQTWEVKPDLVAPGVSIDSTYPRGYLALNGTS 415

RESULT 9
 Q9K66 PRELIMINARY; PRT; 757 AA.
 AC Q9K66; 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE MINOR EXTRACELLULAR SERINE PROTEASE.
 GN BH3763.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Bacillus.
 CC NCBI_TaxID=8665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis."
 RT Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001519; BAB07482.1; .
 DR InterPro: IPR001337; PA.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF02225; PA: 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; 1.
 DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE: PS00138; SUBTILASE_SER; 1.
 KM Protease; Complete proteome.
 SEQUENCE 757 AA; 82516 MW; F6E28CF597435E1 CRC64;

Query Match 8.6%; Score 298.5; DB 2; Length 757;
 Best Local Similarity 21.6%; Pred. No. 5.9e-08;
 Matches 127; Conservative 97; Mismatches 181; Indels 183; Gaps 25;

QY 96 ENPK---KEDKVVYIAEFKDKSEKAIKELSSKNTKLYLYDRIFNGSAIETTPDNL 152
 DB 30 EKPRDLADEYVIVLTNNDLETTIKDYER--SISGELRMTE-RLNGFSFQLEPEEVE 86
 QY 153 KIKQIEGSSVERAOKVOPMMNHAKKEIGVEEAIDYLSINAPFGKNGFGKGVNISNIDT 212
 DB 87 KLTYSYVERVDHVTYETATINESAPFIGAQ---VRQMDDEBVHLGKVKKAVIDT 142
 QY 213 GTDYHKKAMRIDDAKASMRKKEDLKTGNWYLSDKIPAFNYNGCKITVEKYDDGR 272
 DB 143 GIDYTH-----PDQSSYKGGDYFVD-YD-- 166
 QY 273 DYFDP-----HGHIAGILAGNDPEODIKNFNGIDGIAFNAQIFSYKMYSDAG 320
 DB 167 ---DPMETIASGCPPTLHGTHVSGITAA-----GQVGVAPAEITVIRALGREG 214
 QY 321 SGFAGDETFMHAIEDSIKHNVDVSVSGFTGTGLVGEKYMQAIRALKRA--GIPMYVA 377
 DB 215 QGTT--EQVIAIEKAVEGCVINLSGNVYNG---PDMTSLALDAAVEGVAAVTS 268
 QY 378 TGNVYTSASSSSWDLVANNHLMKMTDTGNVTRFAHEDAIYV-ASKAKNOTVEFDKNIGCE 436
 DB 269 NGN---SGPNMW-----TVGSPGTSKKAISVGASAPPLNTPY----- 302
 QY 427 SFKYRNIGAFPEKSKITTNEDGKAP---SKLKVYIYGKGDODLIGDLRGKIAVMDR 492
 DB 303 -----LTFGEENETISLYPSSGGLPMAFKRDLPMIDVGYTEKEMEGVADAGKVVLIK 356
 QY 493 ---IYTKDLKNAFKKAMDGARAIVVNTVNYNRDNTLTPAMGYEADGTSQVFSIS 549
 DB 357 GAVPTEKVMH---AVAKAANGVLIYN-----TPGFFTGW-IEGCVINIVVISIT 402
 QY 550 GDDGVKLMNMINPDKKTEYK---RNKKEDEKDKLEQYYPIIDMESFNKPNVDEKEID 605
 DB 403 REDGEFLLEOLELQKNEKELTIRTKEDF-----VALFSSGCP----- 442
 QY 606 FRFADDTKELYKEDIYPAGSTSMGPRIIDLKLPDVAPGKNIKSTL 653
 DB 443 -----VHTWD-----VKPDVAAPGVSDSTI 464

RESULT 10
 O53637 PRELIMINARY; PRT; 1150 AA.
 AC O53637; 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE SCPB.
 GN SCPB.
 OS Streptococcus agalactiae.

OC Bacteria: Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-78-471;
RA Chmouryulina I., Suvorov A., Cleary P.;
RL Infect. Immun. 64:0-0(0).
DR EMBL: U56908; AAB1762.1; -
DR HSSP: Q45670; IDBI.
DR MEROPS: S08_020; -
DR InterPro: IPR001899; Gram_pos_anchor.
DR InterPro: IPR003137; PA.
DR InterPro: IPR00209; Peptidase_S8.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR Pfam: PF02225; PA; 1.
DR Pfam: PF00082; Peptidase_S8; 4.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; UNKNOWN; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
SEQUENCE 1150 AA; 126309 MW; 7BA9366BEC71633C CRC64;

Query Match 8.6%; Score 296.5; DB 2; Length 1150;
Best Local Similarity 24.3%; Pred. No. 1.4e-07;
Matches 122; Conservative 78; Mismatches 163; Indels 139; Gaps 24;

189 LKSNMPCGNPGRGVINIDGTGRHKAIRIDDAKASRFKEDKGTDKN--- 245
111 VKTLOEKAG---GAGTVAVVIDGFPDKNEAMKLTIDTKARYO-SKEDLEKAKKEGIT 166
246 ---WLSDKIPHAFTYNGKLTVEKYDDGHDYDPHGMHAGILAGNDTODIKNFNGID 302
167 YGEVNDKVAHYDYSDKGTAVDQ-----EHGTHVSGILSGN-ASEKREPERLE 216
303 GIAPNAOIFETK-----YSDAGSGFAGDETMHATDSIKHNVYVSSGFTGGLVG 357
217 GAMPEDALLMREYVINGLADYARNA-----QATRDAINLGAKYINMSFGNAALAYAN 270
358 ---KRYMOAIALKRAKAGIPVAVATGNVATSSASSSMDLVANNHLMPTGVTATRAHEDA 415
271 LPDETAKAFYAKSKGVSIYTSAGNDSSFGCKTRPLA--DH--PDYGVGPAAADST 325
416 IAVA---SAKQVTEFEDKVNIGSGEFYRNIGAFEDKSKITTNEDGTAPSK-LKFVYIGK 472
326 LTVASVSPKQLETETVAVKTAQDQK-----EMPVLTNR--FEPKKAQVAYANR 374
473 GQDODLIGDLRGKIAVMRIYTKDLKNAFKKAMDKGARAIMVYVNTVYNNRD----- 525
375 GTEKEDF-KDYKGRKIALIER-GDIDFKDKIAKAKGAVGLI-----YDNDKGFPIEL 427
526 -NMTELPAWGEADEGCKSQVFSISGDDGVKLMNMNIPDKTEKTKRNNKDKPDKLEQY 584
428 PNDQMPAA-----FISKDGLLKD--NP----- 450
585 PIDMESFNKPNVGEKEIDFKFAPDTDEKLEYKEDIIVAGS-----TSMGPRIDL 637
451 -----OKTTFNATPK-----VLPTAGSTKLSRSSMGLTADGN 484

RESULT 11

09KEM1 PRELIMINARY; PRT; 799 AA.
AC 09KEM1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE MINOR EXTRACELLULAR SERINE PROTEASE (EC 3.4.21.).

GN VPR OR BH0831.
OC Bacillus halodurans.
OC Bacteria: Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RA MEDLINE-20512582; PubMed-11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.,
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
DR EMBL: AF001510; BAB04550.1; -
DR InterPro: IPR003137; PA.
DR InterPro: IPR00209; Peptidase_S8.
DR Pfam: PF02225; PA; 1.
DR Pfam: PF00082; Peptidase_S8; 2.
DR PRINTS: PR00723; SUBTILISIN.
DR PROSITE: PS00136; SUBTILASE_ASP; 1.
DR PROSITE: PS00137; SUBTILASE_HIS; 1.
DR PROSITE: PS00138; SUBTILASE_SER; 1.
KW Protease; Hydrolase; Complete proteome.
SEQUENCE 799 AA; 86192 MW; AF0FF17C2BD34D8A CRC64;

Query Match 8.3%; Score 285.5; DB 2; Length 799;
Best Local Similarity 21.4%; Pred. No. 3.1e-07;
Matches 140; Conservative 96; Mismatches 196; Indels 221; Gaps 30;

63 IDNNTSNEAKIRE-ENS----KSQGDYDSEFNKNTENPKKEDKVVYIAEFKDESGE 117
43 IDTSSALFVYIEIDPSTIEAKHG-----QKSKMELKQARGSVIEQID----- 90
118 KAIKELSLKNTVLYYDRIENGSAIETTDPDLNKTQIEGISSVERACKVOPMNNHAR 177
91 -----LVPSSVTHYEDLFSGFALFPAHQIPSTLIGDYHAV--YPIIEVEVETDP 141
178 KEIGVE-----FAIDYKLSNAP-GKN-----FDGRGMVSNIDGTGRHKAIRID 224
142 DEVAIEEDAYSPEMLD-----SAPFTGANDAMEAGYIGESITVAIILDGVYTH----- 190
225 DDAKASREKEDLKGDKNYWLSDKIPHAFTYNGKLTVEKYDDGGRD--FDP----- 277
191 -----PDLVHAFQYKKWDF-IDNNDPQETPPDPRGIER 225
278 -HGMHAGILAGNDTODIKNFNGIDGIAIPNAQIFSKMTSDAGSGFAGDETMFHAIEDS 336
226 THGTHVAGTVAAAGL-----IKGVAPDANLAVRVLGPGRGSTAG--VIAGIERA 274
337 IKHNVDPVSSGFTGGLGKRYQWAIKRLKAGIPMVAVATGNVATSSASSSMDLVANN 396
275 VQGCADIMNLSLNT-----LNDPDPATST-ALDM-AVAEG 308
397 HLMKTDGNTTTRAHEDAIYASAKNQV-----EFDKVNIGSGEFYRNIGAFEDKSK 451
309 VAVATSNGN-----SGPNMWTGSPGTSDAISVGTATRLPYKKY-----KAS 350
452 ITTNEDGTAPS-----KLKFVYIGKQDODLIGDLRGKIAVMRI 493
351 VFTS-DEIDYPSADIMGFPSEDELELDEGETVEYAFAGLCKPDPDFGVDEGRIALIVRG 409
494 YTKDLKNAFKKADKARALIMVYVNTVYVRDMMTELPAWGEAD-EGTSQVFSISGD 552
410 EIPPEVKA-----ENKKAAGAVGALITYN-----VAGVOPYPVGLAIPIMLSNED 455
553 GVKLMNINPDKRTTEYKRNKEDFKDLQYYPIDMESFNKPNVGEKEIDFKFAPDT 612
456 GLMKRNLENGQNT-----VTFSIEF 476
613 DKLEYKEDIIVAGSISWGPRIID-LKLPVSAAPGKNIKSTLVANINGKSTYGY 664

Db 477 DK-LVGETV---ADFSRRGPMHTWIKRPDVSAPEVAIVSTRIPRQDDPYGY 525

RESULT 12

Q9L908 PRELIMINARY; PRT; 1134 AA.
 ID Q9L908
 AC Q9L908;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE C5A PEPTIDASE (FRAGMENT).
 GN SCPB.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1311;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-130;
 RA Adderson E.E., Bohnsack J.F., Takahashi S.;
 RT "Genetic Polymorphisms of Group B Streptococcus agalactiae scpb."
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF189004; AAF04282.2;
 DR InterPro; IPR003137; PA.
 DR InterPro; IPR00209; Peptidase_S8.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 4.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 FT NON_TER 1134 1134
 SQ SEQUENCE 1134 AA; 124414 MW; 723B8DC8B330B73 CRC64;

Query Match 8.1%; Score 279.5; DB 2; Length 1134;
 Best Local Similarity 24.3%; Pred. No. 1.1e-06;
 Matches 122; Conservative 75; Mismatches 167; Indels 139; Gaps 25;

Qy 188 YLKSNAPFGNFGRCGVINIDGTDRHKAIRIDDAKASMKKEDLKTGDKNY-- 245
 Db 110 HVKTLQEKAGK--GAGTVAAVIDAGFDKNHAWRLDTKTARYQ-SKENLEKAKKEHGI 165
 Qy 246 ---WLSDKIPAHFNYNGKITVEKYDDGRDYFDPHGHIAGILAGNTEODIKNFNGI 301
 Db 166 TYGEMVNDKVAIYHYDYSKDGKNAVDQ-----EGHTHVSGLISGN-ASEMKEPEYRL 215
 Qy 302 DGIAPNAQIFSYKM-----YSDAGSGFAGDETFMFAIEDSIKHNVDVSVSSGFTGTGLV 356
 Db 216 EGAMPEAOQLLMRVEIVNGLADYARNYA-----QAIRDAVNLGAKVIMNSFGNALAYVA 269
 Qy 357 G--EKYWAIRALRKAGIPMVVATGNVATSASSSSMDLVANNHLMKTGTGNTRTAAHED 414
 Db 270 NLPDETKAKFDYAKSGVSIYTSAGNDSSEFGKPRPLPLA--DH--PDYGVVGTTPAAADS 324
 Qy 415 AIAVA--SAKNQTFEEDKNIGESFKYRNIGAFEDSKITTNEDGTAPSK-LKFFVYIG 471
 Db 325 TLTVASYSPOQLTETATVKTDDHDK-----EMPYISTNR--PEPKAIYAYAVAN 373
 Qy 472 KGQDODLIGDLRGKIAVMDRIYTRDKLNAPKAMDKAGARAIMVNTVYVYNRD----- 525
 Db 374 RGTKEEDF-KDVEGKIALIER-GDIDFKDKIANAKKAGAVGLI-----YDNODKGFPIE 426
 Qy 526 ---NMTELPAWGYEADDEGKTSQVFSISGDDGVLLMNMNINPDKTEVKRRNKKEDFKOLEY 583
 Db 427 LPNDOMPAA-----FTSRDGLL-----KD----- 448
 Qy 584 YPIDMESFNKPNVGDKEIDFKFAPDTEKELYKEDIIVPAGS-----TSMGPRIDL 636
 Db 449 -----NPP-----KITTFNATPK-----VLPTASGTKLRSFSSWGLTADG 483
 Qy 637 LKPDVSAPEGKNIKSTLVINGK 659

Db 484 NIKPDIAPGQDILS--SVANNK 504

RESULT 13

Q9K2K8 PRELIMINARY; PRT; 1130 AA.
 ID Q9K2K8
 AC Q9K2K8;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE C5A PEPTIDASE (FRAGMENT).
 GN SCPB.
 OS Streptococcus agalactiae.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
 OC Streptococcus.
 OX NCBI_TaxID=1311;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-125, AND GM;
 RA Adderson E.E., Bohnsack J.F., Takahashi S.;
 RT "Genetic Polymorphisms of Group B Streptococcus agalactiae scpb."
 RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF189003; AAF04281.2;
 DR EMBL; AF189002; AAF04280.2;
 DR InterPro; IPR003137; PA.
 DR InterPro; IPR00209; Peptidase_S8.
 DR Pfam; PF02225; PA; 1.
 DR Pfam; PF00082; Peptidase_S8; 3.
 DR PRINTS; PR00723; SUBTILISIN.
 DR PROSITE; PS00136; SUBTILASE_ASP; 1.
 DR PROSITE; PS00137; SUBTILASE_HIS; UNKNOWN_1.
 DR PROSITE; PS00138; SUBTILASE_SER; 1.
 FT NON_TER 1130 1130
 SQ SEQUENCE 1130 AA; 124043 MW; E04D343FBA12E994 CRC64;

Query Match 7.5%; Score 257.5; DB 2; Length 1130;
 Best Local Similarity 23.5%; Pred. No. 1.5e-05;
 Matches 118; Conservative 75; Mismatches 167; Indels 143; Gaps 25;

Qy 188 YLKSNAPFGNFGRCGVINIDGTDRHKAIRIDDAKASMKKEDLKTGDKNY-- 245
 Db 110 HVKTLQEKAGK--GAGTVAAVIDAGFDKNHAWRLDTKTARYQ-SKENLEKAKKEHGI 165
 Qy 246 ---WLSDKIPAHFNYNGKITVEKYDDGRDYFDPHGHIAGILAGNTEODIKNFNGI 301
 Db 166 TYGEMVNDKVAIYHYDYSKDGKNAVDQ-----EGHTHVSGLISGN-ASEMKEPEYRL 211
 Qy 302 DGIAPNAQIFSYKM-----YSDAGSGFAGDETFMFAIEDSIKHNVDVSVSSGFTGTGLV 356
 Db 212 EGAMPEAOQLLMRVEIVNGLADYARNYA-----QAIRDAVNLGAKVIMNSFGNALAYVA 265
 Qy 357 G--EKYWAIRALRKAGIPMVVATGNVATSASSSSMDLVANNHLMKTGTGNTRTAAHED 414
 Db 266 NLPDETKAKFDYAKSGVSIYTSAGNDSSEFGKPRPLPLA--DH--PDYGVVGTTPAAADS 320
 Qy 415 AIAVA--SAKNQTFEEDKNIGESFKYRNIGAFEDSKITTNEDGTAPSK-LKFFVYIG 471
 Db 321 TLTVASYSPOQLTETATVKTDDHDK-----EMPYISTNR--PEPKAIYAYAVAN 369
 Qy 472 KGQDODLIGDLRGKIAVMDRIYTRDKLNAPKAMDKAGARAIMVNTVYVYNRD----- 525
 Db 370 RGTKEEDF-KDVEGKIALIER-GDIDFKDKIANAKKAGAVGLI-----YDNODKGFPIE 422
 Qy 526 ---NMTELPAWGYEADDEGKTSQVFSISGDDGVLLMNMNINPDKTEVKRRNKKEDFKOLEY 583
 Db 423 LPNDOMPAA-----FTSRDGLL-----KD----- 444
 Qy 584 YPIDMESFNKPNVGDKEIDFKFAPDTEKELYKEDIIVPAGS-----TSMGPRIDL 636
 Db 445 -----NPP-----KITTFNATPK-----VLPTASGTKLRSFSSWGLTADG 479
 Qy 637 LKPDVSAPEGKNIKSTLVINGK 659

DB 480 NIKPDIAAPGODILS--SVANNK 500

RESULT 14

039547 PRELIMINARY; PRT; 731 AA.

AC 039547; 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE PREPRO-CUCUMISIN PRECURSOR.
 OS Cucumis melo (Muskmelon).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid 1; Cucurbitales; Cucurbitaceae; Cucumis.
 NCBI_TaxID=3656;

SEQUENCE FROM N.A.

STRAIN-TERESA; TISSUE-FRUIT;
 MEDLINE=95105149; PubMed=7806492;
 Yamagata H., Masuzawa T., Nagaoaka Y., Ohnishi T., Iwasaki T.;
 "Cucumisin, a serine protease from melon fruits, shares structural
 homology with subtilisin and is generated from a large precursor";
 J. Biol. Chem. 269:32725-32731(1994).
 DR EMBL: D32206; BAA06905.1; .
 DR HSSP: P29599; 1ST3.
 DR MEROPS: S08.092; .
 DR Mendel: 7795; Cucme:1086;7795.
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 1.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00138; SUBTILASE_SER; UNKNOWN_1.
 KM Signal; Serine protease.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 731 PRE-PRO-CUCUMISIN.
 FT CHAIN 111 731 NATIVE CUCUMISIN.
 FT CHAIN 111 615 MATURE AUTOLYZED CUCUMISIN.
 SO SEQUENCE 731 AA; 78819 MW; 93270A9433A366ED CRC64;

Query Match 7.4%; Score 256; DB 10; Length 731;

Best Local Similarity 24.1%; Pred. No. 9.8e-06;
 Matches 148; Conservative 94; Mismatches 211; Indels 162; Gaps 31;

94 NTEPKKEDVVVIA-EFKKESG---EKAIKEL---SLKNTFVLYTDRIENGSALET 146

25 DSDDGKNITVYTMGRKLEPDSALHNRAMLEQVVGSTFAPESVLTHTYKRSFNGFAVKL 84

147 TPNDLRIKQIEGSSVERAQKQVPMNMHARKEIGVEAIDYLSINAPFGKNFDGRGV 206

85 TEEBAEKIAMEGVSV-----FLNE-INELHTTRSMDFL-----GFPVTPRRSQV 130

207 ISNIDG---TVRHKMRIDDDAKASMRFKEDLCTDKNYWLSDKIPHAENYNGSKI 263

131 ESNVVGVLDTGIMPESPDDEGFSPPRKMGCTCTSNFRKRIIGARSYHIRPI 190

264 TVEKYDGRDYFDPHGHINGLAGNDTEODIKNFNCID-----GIAPNOIEFYKMSD 318

191 SPGDVNPGRD-TMGGHTHTASTAGGLVSO--ANLYGLGIGTAGVPLAKRIATAIKYCMW 247

319 AGSGFAGDETFMFAIEDSIKHNVDVSVSSG-----FTGTGLVGEKYYQAIKALRKAG 371

248 DG---CSDTDILAAYDAIDAGVDIISLVGAGNPRHYFVDALIGS--FHAV-----ERG 298

372 IPMWVATG---NATISASSSSMDLVANNHLKMTDGTGNVTRTAHEDALIAVASKNQTVE 427

299 ILASNSAGNGCPNFTTASLSPW-----LLSVAATMDRK 334

428 FDKYNIS-GESEFYRNIGAF-----FDKSKITTNDDGTAKPSKLEFY 469

335 VTVOVQINGSGFQGVSTINTDNOYPIVLSGRDIPNTGFDKSTKSTFCIDKSVNPLLK--- 391

QY 470 IKGQDODLIGDLRGKIAMDRITYTKDLNARKKAMDKARAIWVNYNYNRDNMT 529

DB 392 -----KIVCEASFG---PHEFFKSLDGAAGVLAAMTSNTRYD---AASY-P 430

QY 530 LPAMGYEAD--GTRKQVFSISGDDGVKLMNMINPKKTEFKVKNKNEEDFDKLEQYYPID 587

DB 431 LPSSVDPNDLTLRTYISIR-SPGATIR-----KSTTI-----LNASAPV- 471

QY 588 MESFNSKPNVGDKEIDFKFAPDTKELYEDIIYPAGSTSGPRIDLLK-PDVSAPG 646

DB 472 VSEFSSKGNP-----RATKDVTKPDIS-----GPVEILAAWPSVAAPVG 510

QY 647 KNIKSTL-NVINGKS 660

DB 511 GIRRMTLFNITSGTS 525

RESULT 15

09FII2 PRELIMINARY; PRT; 840 AA.

AC 09FII2; 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE SERINE PROTEINASE.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosid 1; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;

SEQUENCE FROM N.A.

STRAIN-COLUMBIA;
 RC MEDLINE=99397451; PubMed=10470850;
 Kaneko T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
 Miyajima N., Tabata S.;
 "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
 RT Sequence features of the regions of 1,011,550 bp covered by seventeen
 RT P1 and TAC clones";
 RL DNA Res. 6:183-195(1999).
 DR EMBL: AB017065; BAB09160.1; .
 DR InterPro: IPR000209; Peptidase_S8.
 DR Pfam: PF00082; Peptidase_S8; 2.
 DR PRINTS: PR00723; SUBTILISIN.
 DR PROSITE: PS00136; SUBTILASE_ASP; UNKNOWN_1.
 SO SEQUENCE 840 AA; 90616 MW; 0A97E5641FEDDE84 CRC64;

Query Match 7.3%; Score 252.5; DB 10; Length 840;

Best Local Similarity 22.5%; Pred. No. 1.8e-05;
 Matches 150; Conservative 82; Mismatches 233; Indels 201; Gaps 30;

56 VSKNPYIDNNTSNEEKIKE-----ENSNSOG-----DYDSFVNKNTEN 97

DB 41 VTLKQPIVHLFEQDLKHKRSKTPKLRPNRNSRKHGSKIPSVQSHDFLRTKLK- 99

QY 98 PKKEDKVVVIAEFKDKESGKATKELSLKNTKLYLYDDTFNGSAIETTPNDLTKQI 157

DB 100 -----GEKYK-----LYSYHLINGFALFINSQAEEKSMR 131

QY 158 EGISSEVERAQKQVPMNMHARKEIGVEAIDYLSINAPFGKNFDGRGVISNIDTGYDR 217

DB 132 KEVANIYLDVSVTRATYTPQFMGLPGA-WVKE---GGFEIAGEGVITIGIDTIDPN 186

QY 218 HKAMRIDDDAKASMRFKKE-----DLKGTDKNYWLSDKIPHANVNYNGSKITVEKY 268

DB 187 HPSFN-DNDSKRSYPIPKHSGVCEVTPDPFGSCNKKLI-----GARHPAQSAAVTGICF 240

QY 269 DDGRDYFDP-----HGHIINGLAGNDTEODI---KNFNGIDGAPNAQIFSYKMSDAG 320

DB 241 NSSEDIASPFDDGSHGHHTASVAAGNIGVPIVSNHNFVGAASGAPRAAFISYKALKKSF 300

QY 321 SGFAGDETFMFAIEDSIKHNVDVSVSSGFT-----GTGLVGEKYYQAIKALRKAGIPM 374

```

Db 301 GGFAD--VVAIDQAOQDGVILSL--ITPNRKPQGVAFPFNPIDMALLSAVKAGIFV 356
QY 375 VVATGNVA-----TSASSSSW-----DLVANNHLKMTDIGNVTRTAHEDAIYASA 421
Db 357 VOAAAGTGPAPKTMSSFSFPIFTVGASSHDRVSNLTL--GNNVTIPGMGFAIPTDSG 413
QY 422 KNOTVEFDKVNIGSESFYRNIGAFEDSKITTNEDGTRKAPSKLFEVYIGKODODLIGL 481
Db 414 K-----MYKMSAFHALNSTVD-----KDMYVGECDIENFDQ 448
QY 482 D-LRGKIAVMDRIYTKDL--KNAFKKAMD--KGARAIMVYNTVYNNRDMWTELPMAGY 535
Db 449 DRVSGKLILCS--YSARFVLGLSTIKQALDYAKNLSTATGVIFYIDPY-----VLGF 497
QY 536 EADEGTSQVFSISGDDGKLMNMINP-----DKTEVAKRNKEDFKDKLE 581
Db 498 E-----INPTPMDMPGIIIPSVEDSKTLKLYNSSIQRDVYT 534
QY 582 OYYPIDMESF-----NSNKPNGVDEKEIDFKFAPD-----TDKELYKEDIYVA 625
Db 535 K-----EIVSFGAVAAIEGGLNANFSNRAPKYVYISARGPDPEDNSFNDADVLKPNLVAP- 589
QY 626 GSTSWG 631
Db 590 GNSIWG 595

```

Search completed: December 7, 2001, 11:49:32
 Job time: 202 sec